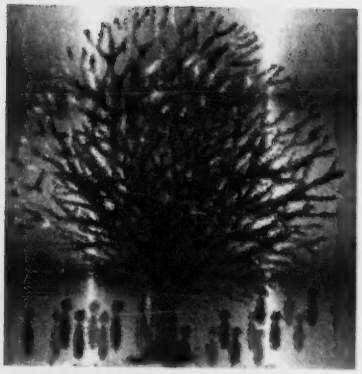
Catalogue no. 91-528-X

Population and Family Estimation Methods at Statistics Canada



Demography Division

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website at www.statcan.gc.ca, e-mail us at infostats@statcan.gc.ca, or telephone us, Monday to Friday from 8:30 a.m. to 4:30 p.m., at the following numbers:

Statistics Canada's National Contact Centre

Toll-free telephone (Canada and United States):

Inquiries line	1-800-263-1136
National telecommunications device for the hearing impaired	1-800-363-7629
Fax line	1-877-287-4369

Local or international calls:

Inquiries line	1-613-951-8116
Fax line	1-613-951-0581

Depository Services Program

Inquiries line	1-800-635-7943
Fax line	1-800-565-7757

To access this product

This product, Catalogue no. 91-528-X, is available free in electronic format. To obtain a single issue, visit our website at www.statcan.gc.ca and browse by "Key resource" > "Publications."

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "About us" > "The agency" > "Providing services to Canadians."

Statistics Canada
Demography Division

Population andFamily Estimation Methods at Statistics Canada

Demographic Analysis Section

Published by authority of the Minister responsible for Statistics Canada

© Minister of Industry, 2012

All rights reserved. Use of this publication is governed by the Statistics Canada Open Licence Agreement (http://www.statcan.gc.ca/reference/copyright-droit-auteur-eng.htm).

March 2012

Catalogue no. 91-528-X

ISBN 978-1-100-19900-9

Frequency: occasional

Ottawa

La version française de cette publication est disponible sur demande.

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued cooperation and goodwill.

Symbols

The following standard symbols are used in Statistics Canada publications:

. not available for any reference period

.. not available for a specific reference period

... not applicable

0 true zero or a value rounded to zero

0^s value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded

^{p} preliminary

^{r} revised

x suppressed to meet the confidentiality requirements of the *Statistics Act*

^{E} use with caution

F too unreliable to be published

Acknowledgements

This manual was prepared by the Demographic Estimates Section, Demography Division under the editorship of Shirley Loh. The written work was based on numerous sources and represents the collaborative work of many individuals in Demography Division and other related divisions, such as Social Survey Methods Division, Income Statistics Division and Health Statistics Division at Statistics Canada. Demographic methods also rely on a number of administrative records maintained by various government bodies responsible for managing these programs; namely, the Canada Revenue Agency and Citizenship and Immigration Canada and the various provincial and territorial bodies responsible for reporting vital events to the bureau.

Inspired from previous editions of this report, the implicit contribution of many others is thus acknowledged. While many people provided valuable input or assistance, special recognition must be given to André Cyr, Vincent Dale, Sylvie DeBlois, Hubert Denis, Hélène Landry, André Lebel and Denis Morissette who spent countless hours editing and providing guidance and information in all stages of the development of the manual. Selected chapters were written by or largely based on the work of the following persons in Demography Division: Julien Bérard-Chagnon, André Cyr, Denis Morissette and Peter Wilkinson. Christian Thibault and Peter Dick (Social Survey Methods Division), Geneviève Ouellet and Julie Dussault (Demography Division) also offered their expertise, background information and editorial comments.

Finally, Nicole Montsion and Janie Saumure provided technical expertise in the publication process.

Table of Contents

Introduction 8

Chapter 1 Postcensal and intercensal population estimates, Canada, provinces and territories 14

Chapter 2 Base population 28

Chapter 3 Births and deaths 46

Chapter 4 Immigration 50

Chapter 5 Net non-permanent residents 52

Chapter 6 Emigrants, net temporary emigrants and returning emigrants 58

Chapter 7 Interprovincial migration 74

Chapter 8 Subprovincial estimates of population 83

Chapter 9 Population estimates by age, sex and marital status 96

Chapter 10 Estimates of census families, economic entities and households 112

Glossary 121

References 130

List of Tables

Table I Availability of population estimates and components of demographic growth 10

Table 1.1 Sources and references of postcensal population estimates — Component data 17

Table 6.1 Calculation of the delay factors: An illustration for production year 2011 64

Table 9.1 Percent distribution of husband-wife families by age group of wife and age group of husband according to the legal marital status concept: Ontario, 2001 Census 105

Table 9.2 Calculation of the incidence of widowhood according to the legal marital status concept: An example for females aged 45 to 49 years, Ontario, 2001/2002 107

Table 10.1 Sources, references and role of data used to produce estimates of census families, economic entities and households 114

Table 10.2 Possible situations for calculating census correction coefficients 115

List of Figures

Figure 1.1 International migration flows for Canada 15

Figure 1.2 Lexis diagram showing intercensal estimation 23

Figure 5.1 Non-permanent residents flows by broad type 53

Figure 9.1 Transition from a distribution of demographic events by age and period to a distribution by age and birth cohort 98

Introduction

Population and family estimation methods at Statistics Canada

Demographic estimates are the cornerstone of statistical measurement for the population and are put to a wide variety of uses to better understand the economic and social situation in Canada. Their importance is underscored by the statutory requirements for the calculation of revenue transfers and cost-sharing programs between the various levels of government. As the National Statistical Agency responsible for these data, among others, timeliness, accuracy and reliability of these data are of the utmost importance.

Every five years, Statistics Canada conducts a national census of population, through which a wide range of demographic data is collected on the Canadian population. The census collects information on citizens, immigrants and non-permanent residents (NPR) residing in Canada. While some countries have a system of continuous population registration from which to obtain basic demographic data on the state and movement of the population for non-census years, it is not the case in Canada.

At a five year interval, the census cannot fill the need for timely data that are required by the various statistical programs in Statistics Canada as well as outside users of data. To fill this gap in the information system, Statistics Canada has a Demographic Estimates Program. By using the most recent census data collected by Statistics Canada and administrative data provided by other government departments, various methodological techniques are used to produce estimates of the Canadian population between censuses[1].

In 1987, Statistics Canada published its first manual detailing the procedures used for estimating population, called *Population Estimation Methods, Canada* (Catalogue no. 91-528). The manual represented a direct response to a long-standing policy at Statistics Canada that requires the methods used to produce the Bureau's statistical information to be open to public scrutiny. The manual was updated in 2003 and 2007 and again herein to document conceptual and methodological changes that have been made to the Demographic Estimates Program.

Data quality indicators and analysis of demographic estimates is important information for users of these estimates and to Statistics Canada in fulfilling its role as being the National Statistical Agency responsible for these data. Ongoing quality indicators are

1. Moreover, Statistics Canada has also developed a Population Projections Program. Based on extrapolations of past trends, these projections reflect informed hypotheses of Canada's demographic future. For more information about the Population Projections Program, please refer to *Population Projections for Canada, Provinces and Territories,* Catalogue no. 91-520-X.

available in the annual and quarterly publications[2]. In addition, the 1987 and 2003 versions of the report included a discussion on the data quality of the various components of demographic growth. In the previous and current versions of the report, the discussion on the quality indicators has been removed. The analysis of these quality indicators and special data quality evaluations will be featured on their own in a companion publication on data quality indicators and analysis.

This manual is intended to be a compendium of the methods and the current procedures used by Statistics Canada to produce and release population and family estimates.

Estimates

Essentially, there are two categories of estimates produced by Statistics Canada: postcensal and intercensal. The first category, postcensal estimates, are produced by using data from the most recent census (adjusted for census net undercoverage (CNU)[3]) and estimates of the components of demographic growth since that last census. These components include births, deaths, immigration, net non-permanent residents, emigration, net temporary emigration, returning emigrants and interprovincial migration. Another component, intraprovincial migration is relevant for estimates within subprovincial areas in Canada. All the elements of the production of population estimates will be discussed in this manual.

Three types of postcensal estimates are produced: preliminary, updated and final estimates, referring to the time frame in which they become available. Preliminary estimates are typically available within three to four months after the reference date while updated estimates are usually available within one year. Final estimates, the most accurate postcensal estimates available, typically take two to three years to complete. The production of three types of estimates is the strategy that best satisfies the commitment of Statistics Canada to balance the timeliness and accuracy dimensions of data quality.

Intercensal estimates are produced every five years and reconcile previous postcensal estimates with the latest census counts adjusted for census net undercoverage. They are generated as soon as census population counts and census net undercoverage become available. This process typically takes two years after census data collection to complete.

Disseminated level of detail for the estimates

Population estimates are disseminated at four geographic levels, including province and territory, census division, census metropolitan area and economic region. Demographic

2. Refer to *Quarterly Demographic Estimates*, Catalogue no. 91-002-X, *Annual Demographic Estimates: Canada, Provinces and Territories*, Catalogue no. 91-215-X and *Annual Demographic Estimates: Subprovincial Areas*, Catalogue no. 91-214-X.

3. Unless otherwise noted, the adjustment for the census net undercoverage (CNU) also includes the incompletely enumerated Indian reserves.

estimates at custom-defined subprovincial levels (e.g., census subdivision, urban centres, health regions, etc.) are possible and can be estimated through cost-recovery special tabulations. For timeliness, and because some components of demographic growth are not available until several months after the reference date, three kinds of postcensal estimates are produced: preliminary postcensal (PP), updated postcensal (PR)[4] and final postcensal (PD)[5]. Intercensal estimates are produced using postcensal estimates and counts from two consecutive censuses adjusted for census net undercoverage. According to the level of geography and type of estimates, different demographic characteristics of the population, including age and sex are produced. Table I shows the level of detail and the reference period for which demographic estimates are disseminated.

Table I
Availability of population estimates and components of demographic growth

Geography	Characteristics	Type of estimate*	Frequency
Canada, provinces and the territories	Total population and components of demographic growth	• Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal	Monthly, quarterly and annual estimates
	Population and components of demographic growth • Age • Sex	• Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal	Annual estimates
	Family • Family size and structure	• Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal	Annual estimates
Census division / census metropolitan area / economic region	Population and components of demographic growth • Age • Sex	• Preliminary postcensal • Updated postcensal • Final postcensal • Intercensal	Annual estimates

* Note that the components are not postcensal or intercensal. They are preliminary, updated and final.

Demographic estimates at custom-defined subprovincial levels can be produced by applying synthetic estimation techniques, under the assumption that the larger geographic area's distribution still holds for lower or alternate geographies; or by regression estimation techniques, under the assumption that the known population totals, usually from the census, still hold.

4. The acronym for updated postcensal estimates is PR as this level of estimates is the revised version of the preliminary postcensal estimates.
5. The acronym for final postcensal estimates is PD due to its French term *postcenitaires définitives*.

Accessibility

Statistics Canada's official release bulletin, *The Daily* (Catalogue no. 11-001-X), delivers the first release of statistical data and publications produced by Statistics Canada, such as the various sets of estimates for a given reference date. It can be accessed electronically by visiting Statistics Canada's official website (www.statcan.gc.ca).

In addition, the most current population estimates and projections, as well as other demographic statistics, are available from Statistics Canada's Canadian Socio-Economic Information Management System (CANSIM)[6]. CANSIM is a database on time series data on a wide variety of social and economic topics made available for public use. It can be accessed on Statistics Canada's website, by CD-ROM (Catalogue no. 10F0007XCB), and on Directory Disc (Catalogue no. 10F0005XCB).

More information about demographic data and products can be obtained by contacting Statistics Canada:

Web: www.statcan.gc.ca

E-mail: infostats@statcan.gc.ca or demography@statcan.gc.ca

Toll-free telephone (Canada and the United States):
1-800-263-1136 - Enquiries line
1-800-363-7629 - National TTY line (teletype machine)
1-877-287-4369 - Fax number

Telephone:
1-613-951-8116
1-613-951-0581 - Fax number

Mail:
Statistics Canada
150 Tunney's Pasture Driveway
Ottawa, Ontario
K1A 0T6

Users of population and family estimates

In Statistics Canada, demographic information is used to calibrate sampling weights of many social and household based survey data, and is a fundamental part of the analytical framework for most statistical programs. Current estimates of Canada's population have other wide ranging applications in the areas of planning and program evaluation in both the public and private sectors. For example, the calculation of revenue transfers and grants under various federal statutory programs, as well as cost-sharing agreements between federal, provincial, territorial and municipal governments are highly dependent on demographic data. Statistics Canada is under statutory obligation to provide the

6. Data retrieved through the www.statcan.gc.ca commercial services are subject to specific conditions of use detailed in the Limited Use Data Products Licence Agreement statement associated with these services.

federal government with annual population figures as well as various economic indicators (e.g., gross domestic product) that have been certified by the Chief Statistician of Canada. These figures are used to determine the amounts payable under various federal-provincial fiscal arrangements, such as the Equalization and Territorial Formula Financing (TFF), the Canada Health Transfer (CHT) and the Canada Social Transfer (CST), according to a per capita funding formula. The federal government distributes billions of dollars annually in federal transfers to provinces and territories using population estimates.[7]

The CHT and CST are federal transfers which support specific policy areas such as health care, post-secondary education, social assistance and social services, early childhood development and child care.

The Equalization and TFF programs provide unconditional transfers to the provinces and territories. Equalization enables less prosperous provincial governments to provide their residents with public services that are reasonably comparable to those in other provinces, at reasonably comparable levels of taxation. TFF provides territorial governments with funding to support public services, in recognition of the higher cost of providing programs and services in the north.

Population is a key variable in allocation formulae used by the federal government in defining its annual spending. As billions of dollars depend directly upon these allocation formulae, timely and accurate estimates are extremely important.

The Government of Canada relies upon population data to also inform some of its most fundamental policies. For example, up-to-date information on the evolving demographic situation in Canada serves useful in informing cabinet on its decisions relating to immigration policy. In the administration of public policy, current population figures also serve to increase the cost-effectiveness of program spending.

Provincial and municipal governments also use demographic estimates for planning social programs. These governments further use estimates to establish cost-sharing agreements on the basis of per capita spending formulae. Among the public service users of Statistics Canada's estimates, are education and public health planners, public administrators responsible for policing, criminal justice, municipal administration, and waste and environmental management.

Demographic estimates are fundamental in the calculation of social and economic indicators, including birth rates, death rates, school enrolment rates, unemployment rates, life expectancy, etc. Population serves directly as the denominator in many of these indicators. Survey researchers, whether at Statistics Canada, academia, or in private polling agencies, must use up-to-date population figures in the planning of survey research and in the calculation of sampling weights.

7. "Federal Support to Provinces and Territories", on the Department of Finance Canada's website (http://www.fin.gc.ca).

The private sector also uses demographic estimates for business planning, marketing research and investment demographics. The estimates also help companies to properly segment and target its market, or make sound investment decisions.

Outline of this manual

The first chapter of this manual presents the overall methods used to produce population estimates and focuses on describing the various levels of estimation: postcensal and intercensal estimates. It introduces the concept of the base population and the factors of demographic growth and their components; from components that lead to "natural increase", that is births and deaths, to international effects like "international migration" (immigrants, emigration, net temporary emigration, returning emigrants and net non-permanent residents), to redistributive effects such as "interprovincial migration".

The individual chapters that follow include discussions on data sources, relevant concepts and methodology for each of these components. Each chapter will describe how estimates are produced, what data are used in the calculations and where they come from. This should give the reader a clearer picture of the strengths and weaknesses of each procedure.

Chapter 2 describes how the base population is determined and adjusted following the most recent census of the population. Chapters 3 to 7 describe the components of demographic growth for Canada, provinces and territories. Chapter 3 focuses on information on births and deaths. Chapters 4 through 6 present the components of international migration (immigration is presented in Chapter 4, net non-permanent residents in Chapter 5, and Chapter 6 deals with emigration, net temporary emigration and returning emigrants). Interprovincial migration is discussed in Chapter 7.

Chapter 8 describes the methods used to produce subprovincial population estimates. Chapter 9 focuses on the population estimates by age, sex and marital status. Finally, Chapter 10 describes the methods for estimating census families, economic entities and households. A glossary of principal terms is included at the end of this manual, in Appendix 1.

Chapter 1

Postcensal and intercensal population estimates, Canada, provinces and territories

This chapter describes the types of method used by Statistics Canada to calculate postcensal and intercensal estimates for the total population and for the population by age and sex, at the provincial and territorial levels. The sources of data used to produce these estimates are also given.

1.1 Postcensal population estimates, Canada, provinces and territories

1.1.1 Definition and calculation of provincial and territorial postcensal estimates of total population

Postcensal population estimates are produced using data from the most recent census (adjusted for census net undercoverage (CNU)[8]) and estimates of the components of demographic growth since that census. The data are forwarded from the Census Day to July 1 by taking into account the components of demographic growth between Census Day and June 30 of the census year. The component method used to produce postcensal estimates is a population accounting system, where modifications are made to the current census population adjusted for CNU or most recent estimate by adding and subtracting the components of demographic growth that occur between July 1 and the reference date of the estimate.

The factors of demographic growth and their components are:
Natural increase

- births
- deaths

International migration

- immigrants
- emigrants
- returning emigrants
- net temporary emigration
- net non-permanent residents

Interprovincial migration

- in-migrants
- out-migrants

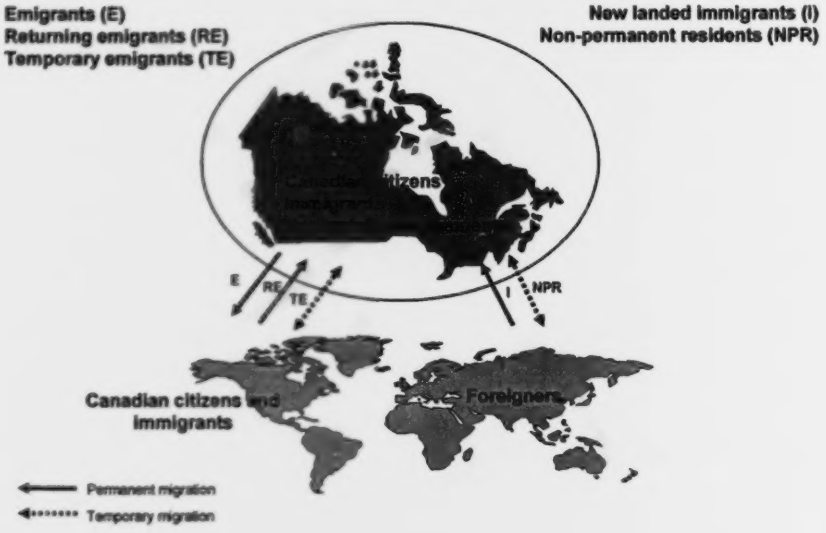
8. Unless otherwise noted, the adjustment for the census net undercoverage (CNU) also includes the incompletely enumerated Indian reserves.

These components can also be divided into two groups, according to the type of data used: those components for which data are readily available, such as births, deaths, and immigration, and those that have to be estimated, such as interprovincial migration, emigrants, returning emigrants, net temporary emigration, and net non-permanent residents (NPRs).

The two components of natural increase, i.e. births and deaths, have similar methodological approach when it comes to estimation. Provincial and territorial Vital Statistics Acts (or equivalent legislation) render compulsory the registration of all live births and deaths within their jurisdictions. Vital statistics universe include births and deaths of all Canadians, immigrants and non-permanent residents (NPR) and exclude foreign residents.

Growth due to international migration represents the movement of population between Canada and a foreign country which involves a change in the usual place of residence.

Figure 1.1
International migration flows for Canada



In the Demographic Estimates Program, international migration consists of five components: immigration, emigration, returning emigrants, net temporary emigration and net non-permanent residents. International migration flows can be categorized as either permanent or temporary. Permanent flows are composed of persons arriving in Canada for permanent residence (landed immigrants), Canadian citizens or immigrants returning to Canada after previously emigrated from Canada (returning emigrants), and Canadian citizens or immigrants leaving Canada to establish a permanent residence in another country (emigrants). Temporary flows relate to foreigners arriving for temporary stay in Canada and leaving after their stay ends (non-permanent residents), as well as Canadian citizens and immigrants living temporarily abroad who have not maintained a usual place of residence in Canada (temporary emigration).

Net non-permanent residents represent the variation in the number of non-permanent residents between two dates, and net temporary emigration represents the variation in the number of temporary emigrants between two dates. Different methodological approaches are used; one for the immigration component, another one for non-permanent residents and a more model based approach for the remaining components of international migration (emigration, returning emigrants, and net temporary emigration).

Finally, the last factor of demographic growth that is discussed is the interprovincial migration. While this factor does not affect the total Canadian population it does affect the provincial and territorial population counts and is a significant challenge for the Demographic Estimates Program.

Table 1.1 shows the sources and references of component data used to generate the postcensal population.

Table 1.1
Sources and references of postcensal population estimates— Component data

Component	Sources and references (if applicable)
Base population	May 16, 2006 Census population adjusted for census net undercoverage and incompletely enumerated Indian reserves. 2006 Census: Statistics Canada, Census of Canada, 2006, Catalogue no. 92-200-XPB. Census net undercoverage: See The Daily, September 29, 2008. Incompletely enumerated Indian reserves: See The Daily, September 29, 2008.
Births and deaths	Statistics Canada, Health Statistics Division. Statistics Canada, Demography Division, Catalogue no. 91-215-X, annual, Catalogue no. 91-002-X, quarterly.
Immigration	Citizenship and Immigration Canada (CIC).
Emigrants	Statistics Canada, Demography Division—from data on emigrant children from the Canada Child Tax Benefit program (CCTB) from Canada Revenue Agency files (CRA) and data from the U.S. Department of Homeland Security, Office of Immigration Statistics.
Returning emigrants	Statistics Canada, Demography Division—based on data from the CCTB program and adjustment factors calculated using CRA files.
Net temporary emigrants	Statistics Canada, Demography Division—based on data from the Reverse Record Check (RRC), 2001 and 2006 Censuses of Canada.
Non-permanent residents	Statistics Canada, Demography Division—based on data provided by Citizenship and Immigration Canada.
Interprovincial migration	Statistics Canada, Demography Division—based on the CCTB program and adjustment factors calculated using CRA files.

Estimates of population are first produced for each province and territory, and then summed to obtain an estimate of the population of Canada.

The component method used in estimating total provincial and territorial populations is expressed as follows:

Equation 1.1:

$$P_{(t+i)} = P_t + B_{(t,t+i)} - D_{(t,t+i)} + I_{(t,t+i)} - (E_{(t,t+i)} + \Delta TE_{(t,t+i)}) + RE_{(t,t+i)} + \Delta NPR_{(t,t+i)} + \Delta N_{(t,t+i)}$$

where for each province and territory:

(t,t+i)	=	interval between times t and t+i;
$P_{(t+i)}$	=	estimate of population at time t+i;
P_t	=	base population at time t (from the census after adjustment for CNU or most recent estimate);
B	=	number of births;
D	=	number of deaths;
I	=	number of immigrants;
E	=	number of emigrants;
ΔTE	=	net temporary emigration;
RE	=	number of returning emigrants;
ΔNPR	=	net non-permanent residents;
ΔN	=	net interprovincial migration.

1.1.2 Provincial and territorial postcensal population estimates by age and sex

Postcensal estimates of the population by age and sex are produced using the cohort component approach, where the population is aged from year to year and the components are organized according to age and sex cohorts. A cohort is a group of persons who experience a certain event in a specified period of time. For the calculation of age and sex estimates, birth cohorts (those persons born during the same year) by sex are used. Therefore the data required for the cohort component method include demographic events, such as deaths, immigration, emigration, that can be directly linked to persons belonging to the same birth cohorts by sex.

Chapter 9 describes the application of the cohort component approach in greater detail. The chapters on the separate components will detail the manner in which the components are organized by age and sex.

1.1.3 Levels of estimates

Producing population estimates between censuses entails the use of data from administrative files or surveys. The quality of population estimates therefore depends on the availability of a number of administrative data files that are provided to Statistics Canada by federal and provincial government departments[9]. Since some components are not available until several months after the reference date, three kinds of postcensal estimates are produced: preliminary postcensal (PP), updated postcensal (PR)[10] and final postcensal (PD)[11]. When all the components are preliminary, the estimate is described as preliminary postcensal. When they are all final, the estimate is referred to as final postcensal. Any other combination of levels is referred as updated postcensal estimates. The difference between the reference date and the release date is three months for preliminary estimates and two to three years for final estimates.

1.2 Intercensal population estimates, Canada, provinces and territories

Intercensal estimates are estimates of population for reference dates between two censuses. They are produced following each census in order to reconcile previous postcensal estimates with the new census counts adjusted for CNU, thus assuring the internal consistency of the estimation system.

The production of intercensal estimates involves two basic steps:

1. the calculation of the error of closure;
2. the linear distribution of the error of closure according to the number of days between intercensal years.

The error of closure is defined as the difference between the postcensal population estimates on Census Day and the population enumerated in that census (after adjustment for CNU). Assuming that the coverage studies that follow each enumeration are unbiased, the adjusted census figures are considered exact.

More specifically, the error of closure is calculated as:

Equation 1.2: $$\varepsilon = \mathbf{P} - \mathbf{P}$$

where

ε = error of closure;

9. In addition to federal and provincial government departments, Statistics Canada also receives data files from the Office of Immigration Statistics, U.S. Department of Homeland Security which is a data source to estimate emigration to the United States.
10. The acronym for updated postcensal estimates is PR as this level of estimates is the revised version of the preliminary postcensal estimates.
11. The acronym for final postcensal estimates is PD due to its French term *postcenitaires définitives*.

P = postcensal population estimate;
P = census population after adjustment for CNU.

The error of closure comes from two sources: measurement errors in any of the components of demographic growth over the intercensal period and errors from the measurement of censal coverage error itself for the current and previous censuses.

The results can be calculated for any disaggregated group, or for any summation of such disaggregation up to and including the total population however disaggregation of the CNU portion has to be modeled as the sample size is not sufficient enough to give reliable disaggregated estimates.

1.2.1 Provincial and territorial intercensal estimates of total population

For the production of intercensal estimates it is assumed that the error of closure is a linear function of the time elapsed since the previous census. The production of intercensal estimates of total population involves two steps: the calculation of the error of closure (ε) as in Equation 1.2, and the distribution of this error uniformly over the intercensal period by an arithmetic function.

Once we have calculated the error of closure we are able to produce the intercensal population estimates for the five years between the two censuses. The intercensal estimates and the residuals are calculated for each month in the intercensal period.

To produce an intercensal estimate of the population at time t we need the following information:

1. The dates of the two censuses (α and β).
2. The date of the estimate that is required (t).
3. The error of closure at the end period (ε_β).
4. The postcensal estimate of the population at time t (P_t).

Intercensal estimates of total population at time t are obtained using the following formula:

Equation 1.3:

$$IP_t = P_t - \left(\frac{t-\alpha}{\beta-\alpha}\right)\varepsilon_\beta$$

Intercensal estimates are then rounded to the nearest integer.

The residual is calculated for each month in the intercensal period. This residual is an added component that is used to balance the adjustments made to the population for the error of closure. It is calculated as follows:

For the month containing the date of the previous census of the intercensal period under consideration (m(α), for example, May 2001):

Equation 1.4:

$$\text{Resid}_{m(\alpha)} = P_{m(\alpha)+1} - IP_{m(\alpha)+1}$$

For the months m(t) between the two censuses m(α) and m(β) (for example, June 2001 to April 2006):

Equation 1.5:

$$\text{Resid}_{m(t)} = P_{m(t)+1} - IP_{m(t)+1} - \sum_{k=m(\alpha)}^{m(t)-1} \text{Resid}_k$$

For the month containing the date of the recent census of the intercensal period under consideration (m(β), for example, May 2006):

Equation 1.6:

$$\text{Resid}_{m(\beta)} = P_{\beta} - CE_{\beta} - \sum_{k=m(\alpha)}^{m(\beta)-1} \text{Resid}_k$$

where

CE = censal estimates.

The sum of all these residuals should equal to the total error of closure.

1.2.2 Provincial and territorial intercensal population estimates by age and sex

The error of closure for each sex and single year of age is the difference between the census estimates (after adjustment for CNU[12]) and the estimated populations, calculated using the same method as is applied to the total population. The production of the intercensal estimates by age and sex involves three steps:

1. the calculation of the error of closure by age and sex;
2. the distribution of this error;
3. a final adjustment to ensure consistency with total population figures estimated independently.

With the exception of ages between 0 and 4 years, and 100 years and over, the error of closure associated with each sex and single year of age is distributed linearly, as a function of the time elapsed since the previous census. Distributing the error of closure

12. The CNU by age and sex was produced by a model-based methodology as reliable estimates were not available due to insufficient sample.

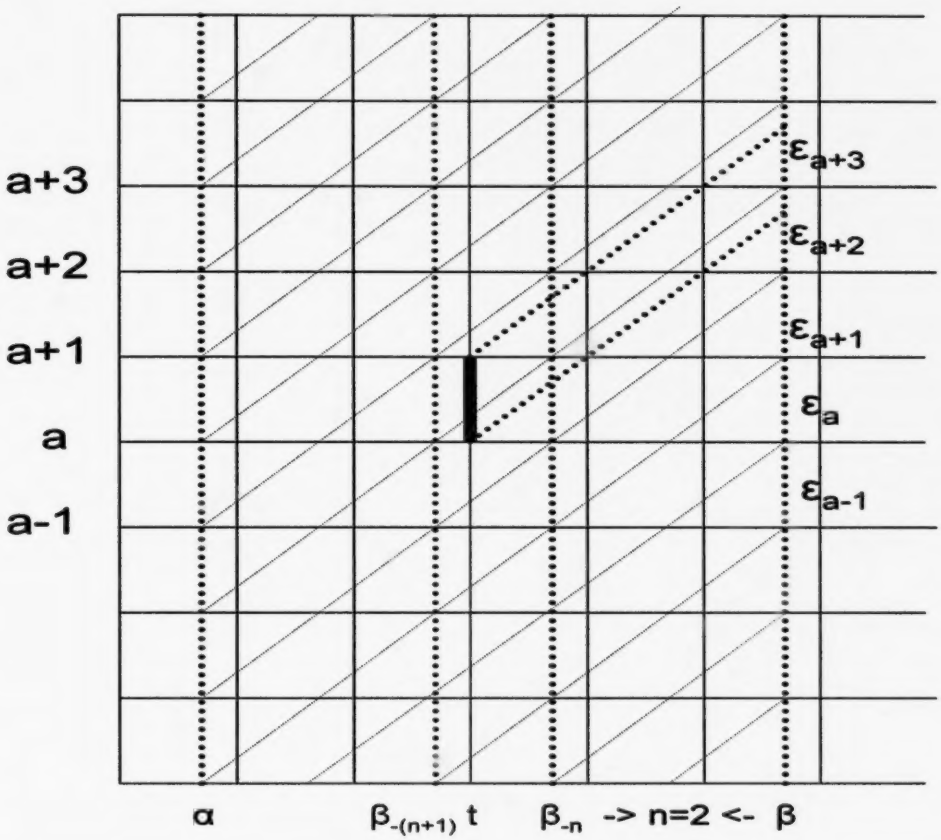
between censuses following specific cohorts generates intercensal estimates. Figure 1.2 shows the method for distributing the error of closure.

To calculate an intercensal estimate at time t for a given province (or territory) p, a particular age a and sex s cohort, we must first define the following:

1. The dates of the two censuses (α and β).
2. The date of the estimate that is required (t).
3. The error of closure by province, age and sex ($\varepsilon_{p,a,s}$).
4. The postcensal estimates of the population at time t for province p, age a and sex s ($P_t(p,a,s)$).
5. The variable n which denotes the number of whole years that separates t and β. For example, if t = 1st of July 2003 and β = 16th May 2006, then n = 2.

The following Lexis diagram (Figure 1.2) is used to illustrate a general example of the intercensal estimate by age and sex.

Figure 1.2
Lexis diagram showing intercensal estimation



The intercensal estimate at time t for province or territory p, for age a and sex s is calculated differently based on the date and age:

A. If $t_a > \alpha$ (meaning that the age cohort or part of the cohort was born after the previous census) the following formula is used:

Equation 1.7:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}}\right) f_{t,a,a+n}\varepsilon(p,a+n,s) + \left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right) f_{t,a,a+n+1}\varepsilon(p,a+n+1,s)\right]$$

where

$f_{t,a,x}$ = is the fraction of the age cohort at time t aged x (which is either a+n or a+n+1 at the time of the current census β), this is the portion of time between α and t in relation to the whole intercensal period.

To calculate this fraction we use:

p = date at the start of births for this cohort;
q = date at the end of births for this cohort.

These are assigned as follows:

If $x = a + n$, then $p = \beta_{-(a+n+1)}$ and $q = t_{-a}$.
If $x = a + n + 1$, then $p = t_{-(a+1)}$ and $q = \beta_{-(a+n+1)}$.

Once we have p and q, we calculate $f_{t,a,x}$ as follows:

Equation 1.8:

$$f_{t,a,x} = \frac{A_{p,q}(\alpha, t)}{A_{p,q}(\alpha, \beta)}$$

where

$A_{p,q}(i, j)$ = area between time i and j of the cohort where the births have occurred between p and q.

It is noteworthy that this area is relative to the size of the cohort (q-p). These results remain valid given that the size of the cohort cancels out in the calculation of $f_{t,a,x}$.

To calculate $A_{p,q}(i, j)$, we need to derive the following variables:

$i' = \max(i, p)$;
$j' = \max(\min(j, q), i)$;
$i'' = \max(i, q)$;
$j'' = \max(j, q)$.

We then calculate the area using the following formula:

Equation 1.9:

$$A_{(p,q)}(i,j) = \left(\frac{(j'-i')(i'-p) + \frac{(j'-i')^2}{2}}{q-p} \right) + (j''-i'')$$

where p=q, then we set $f_{t,a,x} = 1$. This is arbitrary and will not affect the outcome since the condition to have p=q implies that this term of the equation is nil.

B. If (t_a <= α) and (a <= agemax-n-2) meaning no births in the intercensal period are involved and the last age cohort is still bounded by the current census (β):

In this case the general formula described previously can still be used. In fact we can show that if (t_a <= α) and (a <= agemax-n-2), the formula reduces to the following expression:

Equation 1.10:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right]\left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}}\right)\varepsilon(p,a+n,s) + \left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)\varepsilon(p,a+n+1,s)\right]$$

C. If a = agemax-n-1, the age cohorts that will age to the last unbounded age cohort during the intercensal period:

Once the age cohort reaches agemax-n-1, we have to take into account cohorts that are as old or older than the maximum age that is released in the estimates program (agemax) at the time of the recent census (β). At agemax-n-1, we use the error of closure for age max−1 and agemax.

Equation 1.11:

$$IP'_t(p,a,s) = P_t(p,a,s) -$$

$$\left[\frac{t-\alpha}{\beta-\alpha}\right]\left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}}\right)\varepsilon(p,a+n,s) + \left(\frac{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)P_t(p,a,s)}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)P_t(p,a,s) + \sum_{i=a+1}^{agemax} P_t(p,i,s)}\right)\varepsilon(p,age\ max,s)\right]$$

In the case where $\sum_{i=a}^{agemax} P_t(p,i,s) = 0$, we suppose a uniform distribution and the equation reduces to :

Equation 1.12:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right]\left[\left(\frac{t-\beta_{-(n+1)}}{\beta_{-n}-\beta_{-(n+1)}}\right)\varepsilon(p,a+n,s) + \left(\frac{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)+n+1}\right)\varepsilon(p,\text{age max},s)\right]$$

D. If a >= agemax – n these are the remaining cohorts in the unbounded category:

In this last case, we are looking at the age cohorts that are at agemax or older at time β.

Equation 1.13:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right]\left(\frac{P_t(p,a,s)}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)P_t(p,\text{age max}-n-1,s) + \sum_{i=agemax-n}^{agemax} P_t(p,i,s)}\right)\varepsilon(p,\text{age max},s)$$

In the case where $\sum_{i=agemax-n-1}^{agemax} P_t(p,i,s) = 0$, we suppose a uniform distribution and the equation reduces to:

Equation 1.14:

$$IP'_t(p,a,s) = P_t(p,a,s) - \left[\frac{t-\alpha}{\beta-\alpha}\right]\left(\frac{1}{\left(\frac{\beta_{-n}-t}{\beta_{-n}-\beta_{-(n+1)}}\right)+n+1}\right)\varepsilon(p,\text{age max},s)$$

Special adjustment of the intercensal estimate to maintain coherence with the intercensal population estimated by province

Since the error of closure is estimated by cohorts, the intercensal estimates by age and sex will not exactly match the total by province as measured in the first part of this chapter. A final adjustment is done to ensure that both estimates are consistent.

Equation 1.15:

$$IP_t(p,a,s) = \left(\frac{IP_t^{total}(p)}{\sum_{i,j} IP'_t(p,i,j)} \right) IP'_t(p,a,s)$$

where $IP_t^{total}(p)$ is the intercensal estimate measured for province p using Equation 1.3 to Equation 1.6 at time t.

The special case where the intercensal population estimate becomes negative

It can happen, although rarely, that for certain age and sex cohorts for certain provinces or territories that have very low counts, are assigned a negative population count with the above mentioned methodology. In these cases, the counts will be set to zero and the difference will be redistributed proportionately in all the other cohorts. The estimates are then rounded to the nearest integer.

Chapter 2

Base population

A **base population** is the population at the beginning of a period used as a reference or starting point for the estimation process. For postcensal estimates, the base population is the population enumerated in the most recent census, adjusted for census net undercoverage (also referred to as the "censal estimate"[13]). Postcensal estimates as of July 1 of a census year are obtained by the component method, using the most recent census of population adjusted for census net undercoverage (CNU)[14] and taking into account the demographic events that occurred between the Census Day and June 30. The base population for intercensal estimates is referred to as the "postcensal-estimated population".

Since census net undercoverage (census undercoverage minus census overcoverage) is an important aspect of estimating population counts used in the Demographic Estimates Program, this chapter focuses on the census population, adjusted for net census undercoverage. It begins with a brief description of the census collection procedures and the definition of the population universe of the 2006 Census, followed by a discussion on studies used to provide estimates of census coverage error with a section on adjustments for non-enumerated Indian reserves and settlements, and concludes with procedures used for estimating census net undercoverage.

2.1 Censal estimates as the base population

The census requires the participation of the entire population of Canada. Every five years, Statistics Canada conducts a census. The Census of Canada up to and including the 1966 Census has been conducted by interview. Starting from the 1971 Census, two collection methods have been used: self-enumeration and interview. In 2006, about 98% of households were enumerated using self-enumeration. These households either received their questionnaire by Canada Post delivery (Canada Post delivery was used only in 2006) or from a census enumerator. About 2% of households were enumerated using the canvasser method. An enumerator visited a household and completed a questionnaire for the household by personal interview. This method was normally used in remote and northern areas of the country and on most Indian reserves. It was also used in large urban downtown areas where many residents are transient[15].

13. The reference date for the censal estimate is the same as the corresponding census.
14. Unless otherwise noted, the adjustment for the census net undercoverage (CNU) also includes the incompletely enumerated Indian reserves.
15. The respondent was given either a 2A or 2B census questionnaire to complete and help was given if asked for.

For the first time, the 2006 Census offered all households in Canada the option of completing their questionnaire online. Each paper questionnaire had a unique internet access code printed on the front along with the 2006 Census website address. Respondents needed this access code to complete their questionnaire online; the information was directly submitted into the Data Processing Centre system and was verified for completeness. Approximately 18% of households responded via the Internet. Details about 2006 Census data collection and data processing procedures are described in *2006 Census Technical Report: Coverage*[16].

The base populations in the Demographic Estimates Program are derived from the quinquennial censuses between 1971 and 2006. The population universe of the 2006 Census includes the following groups:

- Canadian citizens (by birth or by naturalization) and landed immigrants with a usual place of residence in Canada;
- Canadian citizens (by birth or by naturalization) and landed immigrants who are abroad, either on a military base or attached to a diplomatic mission;
- Canadian citizens (by birth or by naturalization) and landed immigrants at sea or in port aboard merchant vessels under Canadian registry;
- Non-permanent residents:
  - persons with a usual place of residence in Canada who are claiming refugee status on Census Day and members of their families living with them;
  - persons with a usual place of residence in Canada who hold study permits (covering Census Day) and members of their families living with them;
  - persons with a usual place of residence in Canada who hold work permits (covering Census Day) and members of their families living with them.

The population universe of the 2006 Census does not include foreign residents but, since 1991, non-permanent residents are included in the population universe.

Foreign residents have not been enumerated since the 1991 Census. Foreign residents are persons who belong to the following groups:

- government representatives of another country attached to the embassy, high commission, or other diplomatic body of that country in Canada, and members of their families living with them;
- members of the Armed Forces of another country who are stationed in Canada, and members of their families living with them;
- residents of another country visiting Canada temporarily (for example, a foreign visitor on vacation or on business, with or without a visitor's permit).

The definition of the population universe indicates which persons should be included in the census, but not where these persons should be enumerated. The Canadian census uses the modified *de jure* method of enumeration, whereby persons are to be enumerated at

16. See Statistics Canada 2010c.

their usual place of residence, even if they are temporarily away at the time of the census. Persons away from their usual place of residence and residing elsewhere in Canada are to be enumerated at their usual place of residence and are considered temporarily residing at the other location. Persons without a usual place of residence are to be enumerated wherever they happen to be on Census Day.

Each base population for the Demographic Estimates Program (P_t, where t = the census year) is adjusted as follows (unless otherwise noted, adjustments to the base population apply to provincial, territorial and subprovincial levels):

- adjustment of the population for census net undercoverage (CNU);
- addition of independent estimates for incompletely enumerated Indian reserves in 1991, 1996, 2001 and 2006;
- adjustment for early enumeration in 1991 and 1996 in parts of northern Quebec, Newfoundland and Labrador, Yukon and the Northwest Territories;
- addition of estimates of non-permanent residents in 1971, 1976, 1981 and 1986. Since 1991, non-permanent residents are included in the census universe;
- at the provincial level, the first postcensal population estimate is July 1 of the census year. This is obtained by addition or subtraction of the components of growth between Census Day and June 30. At the subprovincial level, the estimate of the July 1 population estimate is obtained by applying to the annual components of growth, a fraction of the year that corresponds to the period between Census Day and June 30. These are adjusted to the provincial/territorial components.

2.2 Adjustment for census net undercoverage (CNU)

Coverage errors are defined as errors caused by the miscounting of the population on Census Day. There are two types of coverage error. **Population undercoverage** refers to the error of excluding someone who should have been enumerated. **Population overcoverage** refers to the error of either enumerating someone more than once or including someone who should not have been enumerated. The latter error is considered negligible. Undercoverage is more common than overcoverage. The net impact of undercoverage and overcoverage on the size of a population of interest is **census net undercoverage** (CNU). Census net undercoverage is calculated as the number of persons excluded who should have been enumerated (undercoverage) less the number of excess enumerations of persons enumerated more than once (overcoverage). Coverage errors are one of the most important types of error since they affect not only the accuracy of the counts of the various census universes, but also the accuracy of all of the census data describing the characteristics of these universes.

Following each census, Statistics Canada undertakes coverage studies to measure coverage errors. Coverage studies provide undercoverage estimates for the 1991, 1996, 2001 and 2006 Censuses at the provincial and territorial levels, and for the 1971, 1976, 1981 and 1986 Censuses at the provincial level only. Estimates of overcoverage at the

provincial and territorial levels are available only for the last four censuses (1991, 1996, 2001 and 2006). Overcoverage for previous censuses was estimated by assuming that the overcoverage-to-undercoverage ratio for each census between 1971 and 1986 was the same as in 1991. The CNU for Yukon and the Northwest Territories prior to 1991 was estimated by assuming that the ratio between the CNU for each territory and the 10 provinces for each census between 1971 and 1986 was the same as in 1991.

For consistency, 1991 Census undercoverage and overcoverage were revised in 1998 to take into account the methodological improvements made in the 1996 Census coverage studies. This revision altered the CNU in all censuses between 1971 and 1986. Similarly, 1996 Census undercoverage and overcoverage were revised in 2003.

The following discussions on the procedures to estimate CNU are based on the 2006 Census coverage studies.

2.2.1 Census coverage studies

Census coverage error of the 2006 Census is measured by three studies. The 2006 Dwelling Classification Survey (DCS) addressed coverage error resulting from dwelling occupancy classification error. Census data were adjusted for this type of coverage error. The 2006 Reverse Record Check (RRC) measured population undercoverage. The 2006 Census Overcoverage Study (COS) measured population overcoverage. Census data are not adjusted for the population coverage error measured by the RRC and the COS. Rather, estimates of census net undercoverage are used in the production of Statistics Canada's demographic estimates of population.

The 2006 Census coverage studies differ from the 2001 Census coverage studies in the following ways:

- The 2006 Census was the first time that the names of persons listed on all of the census forms were available in electronic format. This change greatly increased the efficiency of coverage studies since matching could include the name and not be restricted to demographic characteristics.
- A new coverage study, the COS, was designed to exploit the use of an individual's name for identifying overcoverage. The COS was able to evaluate overcoverage resulting from persons being enumerated more than once with a high degree of accuracy.
- The measurement of population overcoverage was dropped from the RRC. Consequently, much less field collection was required since only those persons that could not be easily found on the census database were sent to the field.
- There is a change in terminology. What used to be called 'gross undercoverage' is now 'undercoverage.' The more complete label is '2006 Census population undercoverage.'
- The Automated Match Study (AMS) was carried out for the 2001 Census and has been in place as a coverage study since the 1991 Census. The AMS was

repeated for the 2006 Census but the results were primarily used for evaluating the COS.

The methodology of each of the 2006 coverage studies is described below.

A. Dwelling Classification Survey (DCS)

One of the potential sources of error in a census is the misclassification of dwellings. When a questionnaire is not returned from a household, the enumerator has to determine if the dwelling is occupied or not. Two types of errors can occur. First, an occupied dwelling can be incorrectly classified as unoccupied. This classification error results in census dwelling and population undercoverage because the dwelling is excluded from the census database. Second, an unoccupied dwelling can be incorrectly classified as occupied. When this error occurs, no questionnaire will be received for this dwelling and it will be subject to non-response follow-up (NRFU). The dwelling will be considered as a non-respondent dwelling and therefore subject to imputation. This would add persons to the census database when, in fact, no one is living at that dwelling thus resulting in population overcoverage. Estimates from the DCS are used to adjust census data for both of these coverage errors.

An additional type of dwelling classification error measured by the DCS is the error incurred when marginal dwellings or dwellings under construction are classified in error as dwellings. This misclassification of dwellings can result in dwelling overcoverage however; census data are not adjusted for these dwellings so census estimates of the housing stock include some degree of overcoverage.

The DCS target population was all non-response private dwellings and all unoccupied private dwellings excluding dwellings in collective collection units (CU), canvasser CUs and Indian reserves CUs. All private dwellings in the sampled CUs that were classified as unoccupied on Census Day or classified as occupied but for which no census form had been returned, were to be checked again in late June or early July 2006 to determine the true occupancy status of the dwellings on Census Day. A DCS questionnaire was used for this purpose.

For each dwelling in the DCS sample found to have been occupied on Census Day, the DCS questionnaire was consulted to determine whether another address was listed where the household members might have been enumerated. If they were found to have been enumerated elsewhere, they were considered as already having been enumerated and therefore they need not add to overcoverage by being included again. The dwelling itself, however was added to the occupied dwelling count.

At this point in processing, the unoccupied dwellings and the non-response dwellings in the sample were separated and the classification of these dwellings was confirmed against final census listing. The questionnaires completed for each sampled CU were matched to the final census listing of unoccupied dwellings. If a match could not be found, the sampled dwelling was discarded and no further processing was required. Dwellings

listed as unoccupied on the census list for which no DCS questionnaire was received were considered as total non-response and went onto the next step of processing. Similarly, the final census listing of all dwellings for which a census questionnaire was not received was used to establish which of the DCS dwellings for which a DCS questionnaire was not received would be considered as total non-response.

Total non-response was addressed by a weighting adjustment while item imputation was used for item non-response. The procedure was the same for the unoccupied dwellings and non-response dwellings. When there was no information for a dwelling, the design weights of the respondents were adjusted to account for the design weight of the non-respondents.

Once the DCS estimates were produced, census data were adjusted for non-response dwellings and for occupied dwellings classified in error as unoccupied. This process resulted in all private dwellings on the database being classified as either occupied or unoccupied. A second procedure was used to impute the household dwelling size and other variables for the selected non-response dwelling. Household size was determined by randomly selecting a dwelling from all dwellings that had completed a census questionnaire in the same CU (nearest neighbour imputation). The complete record from this donor household was then assigned to the non-response dwelling. If no donor was found, then only a household size was assigned.

B. Reverse Record Check (RRC)

The Reverse Record Check (RRC) is a postcensal study carried out to estimate 2006 Census population undercoverage. The target population, which consisted of all persons who should have been enumerated in the 2006 Census, was formed from six sources (sampling frames). The first five frames were used to estimate undercoverage in the ten provinces, whereas estimates for the three territories were calculated based on samples from the last frame only. The six sampling frames of the 2006 RRC are:

1. 2001 Census: all persons enumerated in the 2001 Census;
2. Missed: all persons from the 2001 RRC sample who were classified as missed;
3. Births: all children born between May 15, 2001 and May 15, 2006;
4. Immigrants: all landed immigrants who arrived in Canada between May 15, 2001 and May 15, 2006;
5. Non-permanent residents: all persons from another country, who held employment or student permits, covering May 16, 2006 and persons claiming refugee status on May 16, 2006. Family members living with them in Canada are also in this frame;
6. Health care files: all persons listed in the health care files of Yukon[17], the Northwest Territories, and Nunavut who were living in these territories on May 16, 2006.

17. Some persons from other sources were added.

A problem that exists with using multiple frames is the possibility that persons may be listed on more than one frame. For example, a person in the immigrants frame may have been in Canada on a work permit in May 2001, and thus have been enumerable in the 2001 Census. The person would then be in both the immigrants frame and the census frame if he or she was enumerated, or in the immigrants frame and the missed frame if not enumerated. All potential cases of frame overlap must be identified to avoid double-counting.

Another difficulty is that none of the first five sampling frames covered people who had emigrated, or who were outside the country at the time of the 2001 Census and had returned during the intercensal period. Coverage error estimates do not include these populations.

Sampling fractions were not the same in all strata. To make the sample design more efficient, higher sampling rates were applied in subgroups for which high undercoverage or a lower tracing rate was expected.

The methodology for the territories was changed in 2006. As with previous RRCs, the sampling frames of the three territories were created from their respective health care files. Some files from other sources were added for Yukon in order to improve basic coverage. The people listed in the sampling frames of each territory were then matched by name, sex and age with the 2006 Census response database using exact matching. A manual verification was also performed. Matched people were classified as enumerated, and given a weight of 1. People not classified as enumerated were then stratified by age and sex.

After sample selection and checking the sample for quality of information for different variables of interest (i.e., geographic or demographic), the sample was ready for process and classification. The goal of processing is to determine whether each selected person (SP) was part of the census target population and, if so, to determine whether each SP was enumerated. In addition, processing is undertaken to provide further information for the non-response adjustment.

Most of the work in processing involved searching the RRC version of the 2006 Census Response Database (RRC RDB) to determine whether the SP was enumerated at one of the addresses associated with him or her. The addresses were obtained from various sources including:

- the sampling frame for the selection address;
- updates from tax records;
- the computed-assisted telephone interview (CATI) and paper questionnaires;
- matches with the RRC Response Data Base (RDB) using birth date and sex of the SP and members of his or her household, or, the SP's name, postal code or telephone number.

Two outcomes could result from this process. First, when the SP was found, the classification of 'enumerated' was usually assigned and no further processing was required. An exception was SPs who were later identified as deceased before the census from vital statistics for deaths. Second, when the SP was not found, the case was sent for collection. While collection was taking place, searching the RRC RDB continued. When data from the CATI interview was available, it could be determined whether or not each SP was part of the census target population. If so, the CATI data could enable further searching.

Processing provides the information required to determine which SPs were:

a. 'listed';
b. 'mobile';
c. included in the 'census target population';
d. 'enumerated';
e. 'missed'.

Selected persons for whom one or more of the above-mentioned characteristics could not be determined were considered as non-respondents. Selected persons, who were in the census target population but were not enumerated, thus classified as missed, were the basis for the estimate of undercoverage.

The final weights of the selected persons (SP) began with their initial (or design) weights. The initial weight of an SP from the missed frame was the final weight assigned to him or her during the previous Reverse Record Check (RRC) when the SP was classified as missed. For the other sampling frames, the initial weights were generally equal to the inverse of the probability of selection. The exception was the non-permanent residents frame where the initial weight was higher to account for the small number of non-permanent residents who were not in the sampling frame when the sample was selected. Final non-permanent resident counts were only available after the sample was selected. Initial weights were adjusted to add to these counts.

In order to reduce bias, the initial weights of the respondents had to be adjusted to account for non-response. The weight of the non-respondents was redistributed among the respondents. Where possible, this was done by ensuring that the weight of non-respondents with certain characteristics was redistributed only to respondents with the same characteristics. When a respondent with the same characteristics as a non-respondent could not be identified in a stratum, the stratum was grouped with another stratum deemed similar.

After adjusting for non-response, the estimated number of enumerated persons in the territories has traditionally been lower than the comparable census count. This is likely due to undercoverage of the census target population in the health care files. To address this bias, the weight of SPs selected in a territory was adjusted so that the estimated number of enumerated persons equaled the comparable census count for that territory.

The RRC RDB differs from the final census database in that it does not include imputations made during whole household imputation (WHI), enumerations with an invalid or missing name or an incomplete or invalid birth date, or enumerations added after the start of the RRC data processing phase. People from the target population who are not in the RRC RDB are classified as missed. Census population undercoverage is estimated by the number (weighted) of missed persons less the number of persons excluded from the RRC RDB.

C. Census Overcoverage Study (COS)

Population overcoverage is the number of enumerations in excess of persons who are included in census tabulations more than once, usually twice. This is an error resulting in bias for census counts and estimates because they should only have been included once. Following the 2001 Census, the level of overcoverage due to duplication of individuals was measured by three studies, each one covering a part of the overcoverage: the Automated Match Study (AMS), the Collective Dwelling Study (CDS) and the Reverse Record Check (RRC). The introduction of names to the 2006 Census Response Database (RDB) provides an opportunity to use name matching to measure overcoverage and therefore estimate overcoverage with a single study, the Census Overcoverage Study (COS). The COS is based on a series of automated exact and probabilistic matching operations and manual work. These matching operations also involve the use of various administrative data files. Therefore, the 2006 RRC measures just undercoverage and the CDS is no longer conducted as collective dwellings are covered by the COS.

In principle, the RDB could have been matched to itself to detect duplicate enumerations. However, on a practical level, and for methodological considerations, the COS was conducted in two steps as outlined below.

Step 1—exact matching with administrative data

The first step was based on exact matching procedures, and involved matching the RDB with a set of administrative data files representing a large portion of the census target population. It was expected that this process would directly identify cases of overcoverage. In particular, RDB records assigned to the same administrative record through 'many-to-one' matches were declared to be cases of overcoverage without further review, since they pointed to the same individual from the administrative data files.

The following administrative data files were used:

- 2005 income tax records, supplemented with additional records for taxation years 2000 to 2004;
- Birth files for Canadian citizens born between 1985 and 2003;

- Immigration files for immigrants born outside Canada between 1985 and 2003, to cover children of immigrants not present in the birth files of Canadian citizens born between 1985 and 2003;
- Immigration files for immigrants who arrived in Canada between 2004 and May 16, 2006 (Census Day), given that they would not be on the income tax file for 2005;
- Non-permanent residents files;
- Health care files from Yukon, the Northwest Territories and Nunavut.

As a variety of administrative data files were used, every effort was made to remove duplicates, so the first step exact match would be effective. As the goal of the exact match was to identify each individual in the RDB, it was necessary to only consider cases where a high degree of certainty was achieved. The variables used for this process were name, sex and date of birth.

Overcoverage was identified when two or more RDB records matched to the same administrative record. For evaluation purposes, a sample of these overcoverage matches was manually verified, as well as a sample of the one-to-one cases. An adjustment to the estimate of overcoverage, based on the results of the verification sample, was done to account for false matches whereby two or three records had the same administrative record but did not represent the same individual.

A record in the RDB may have been a match for more than one administrative record, and vice-versa, thus creating a many-to-many match. When two RDB records matched to two administrative records, it was assumed that this grouping contained two valid one-to-one matches. However, a sample of the two-to-two matches was taken to verify this assumption. This review also resulted in cases considered as overcoverage being weighted up and added to the total estimate of overcoverage. All other combinations of many-to-many matches were manually verified and either classified as overcoverage or not, resulting in all of the many-to-many matches being resolved.

In Step 1, for technical reasons, RDB records for the provinces were matched to provincial administrative records, and RDB records for the territories were matched to the records in the territorial administrative Health Care Files. Hence, cases of overcoverage between the provinces and the territories were missed at Step 1, but they were included in Step 2.

Before Step 2, the RDB was split into two parts. Part A consisted of all RDB records that were matched to at least one administrative record, whether overcovered or not. Part B consisted of all RDB records that were not matched to an administrative record, as well as territorial records. The latter was done to take into account provincial-territorial matches that were missed in Step 1.

Step 2—probabilistic match with the Census Response Database (RDB)

Step 2 of the COS is a probabilistic record linkage between RDB records that were not matched with an administrative record (Part B) and the complete RDB (Part A + Part B). Statistics Canada's Generalized Record Linkage System (GRLS) was used for this step.

Within the framework of GRLS, variables such as first name, last name, sex, date of birth and some variables related to geography were considered during the record linkage. GRLS provided results in pairs of individuals with an associated weight that indicates the strength of the match. The higher the matching weight is, the more likely the pair is a good match, thus resulting in overcoverage.

The standard Fellegi-Sunter (1969) approach was implemented in the GRLS. An upper threshold, S2, was established, above which matches were accepted as overcoverage without verification. A lower threshold, S1, below which matches were rejected without further review (i.e., no overcoverage), was also determined to minimize cases of overcoverage below threshold S1. In order to verify cases in the middle zone (i.e., pairs whose matching weight was between S1 and S2), a sample of these matches were selected for manual verification.

In 2006, overcoverage was measured primarily by the Census Overcoverage Study (COS). The total overcoverage estimate comprised individuals overcovered in Step 1, and those deemed overcovered during the probabilistic matching in Step 2. Individuals deemed overcovered in Step 2 whose matching weight was above the upper threshold S2, had a weight of 1. The weight of overcoverage cases identified from the sample between the lower threshold S1 and the upper threshold S2 was determined by the sample design.

To evaluate the COS, the Automated Match Study (AMS) was repeated in 2006. The COS estimates were compared to those of the AMS. The comparison revealed a bias in the COS estimates whereby some pairs identified in the AMS were not found in the COS frames. Since the AMS provided an estimate of overcoverage not included in the COS, the last step in estimating overcoverage was to account for this bias by using the AMS estimates to adjust the COS estimates.

2.3 Calculating census net undercoverage

Let T represent the total or 'true' number of persons in the census target population. Then, let C be the published census count of the number of persons in the census target population. The error in using C instead of T as denoted as N, census net coverage error is defined as:

Equation 2.1: $$N = T - C$$

The censal population P is defined as:

Equation 2.2: $$P = C + N$$

Let U denote population undercoverage. U is the number of persons not included in C who should have been.

Let O denote population overcoverage where O is the number of persons included in C who should not have been. There are two components to O. The first is persons who were enumerated more than once. These duplicate enumerations should not have been included in C. The census coverage studies focus on duplicate enumerations. The second component of O is persons who were included in C who are not in the census target population. Foreign residents visiting Canada, for example, who are listed on a census form as usual residents of a dwelling should not be included in C. Fictitious persons are another example. The number of persons included that are not in the census target population has been seen by previous studies to be negligibly small. Therefore, the 2006 Census coverage studies did not measure this component of coverage error.

Since U refers to persons who should be included in C and O refers to persons who should not be included in C, the difference between T and C is U less O. That is:

Equation 2.3: $$N = U - O$$

The true number of persons in the census target population is then:

Equation 2.4: $$T = C + N = C + U - O$$

An estimate of T is given by $\hat{T}$ where:

Equation 2.5: $$\hat{T} = C + \hat{N} = C + \hat{U} - \hat{O}$$

$\hat{U}$ is an estimate of the number of persons not included in C that should have been; and $\hat{O}$ is an estimate of the number of persons included in C who should not have been. Let us assume that overcoverage from persons included in C who are not in the census target population is zero. Therefore, $\hat{O}$ is restricted to an estimate of the number of duplicate enumerations. It is the goal of the census coverage studies to produce $\hat{U}$ and $\hat{O}$.

Census population coverage error can be usefully expressed as rates relative to the true population: the undercoverage rate R_U is U expressed as a percentage of T. The overcoverage rate R_O is O expressed as a percentage of T. The census net undercoverage rate R_N is the difference between U and O expressed as a percentage of the census target population. These three rates can be estimated by $\hat{R}_U$, $\hat{R}_O$ and $\hat{R}_N$ as follows:

Equation 2.6:

$$\hat{R}_U = 100 \times \frac{\hat{U}}{\hat{T}} = 100 \times \frac{\hat{U}}{C + \hat{N}}$$

Equation 2.7:

$$\hat{R}_O = 100 \times \frac{\hat{O}}{\hat{T}} = 100 \times \frac{\hat{O}}{C + \hat{N}}$$

Equation 2.8:

$$\hat{R}_N = 100 \times \frac{\hat{N}}{\hat{T}} = 100 \times \left[\frac{\hat{U} - \hat{O}}{C + \hat{N}} \right]$$

A positive census net undercoverage rate indicates that undercoverage is larger than overcoverage. That is, there are more people not included in the published census count C than the number of duplicated enumerations. This has been, and continues to be, the experience of the Canadian census. For some domains of interest, however, negative census net undercoverage has recently been observed.

2.4 Adjustments for non-enumerated Indian reserves and settlements

Enumeration is sometimes not permitted on some Indian reserves and settlements or it is interrupted before it can be completed. These areas, a total of twenty-two in the 2006 Census, are called "incompletely enumerated Indian reserves and Indian settlements". Census data for these areas are not available and therefore have not been included in any census tabulations.

Neither the 2006 Census nor the Reverse Record Check is in a position to produce an estimate of the population living in the twenty-two incompletely enumerated Indian reserves and settlements. In order to produce official estimates of population, a model-based methodology was used to prepare estimates of population for these geographical areas.

A two step model was developed to estimate the population. The first step uses a simple linear regression to predict the census count in 2006. The linear regression was constructed using all Indian reserves that were completely enumerated in both the 2001 and the 2006 Census. The model assumes a linear growth from 2001 to 2006 for all provinces with separate estimates, for the intercept and the regression parameters for each province. For each incompletely enumerated reserve, the input variable for the regression

model was either the actual census count in 2001 or the best predicted census count from the 2001 model. The output of the model was the estimated census count in 2006.

The second step is done to produce consistency with the results of the census coverage studies. An adjustment was made to the estimated 'census' count to account for census net undercoverage of all subjected census counts. Census net undercoverage for the incompletely enumerated reserves was estimated by calculating the census net undercoverage rate for all completely enumerated reserves in each province and then applying that rate to the estimated 'census' count of all the incompletely enumerated Indian reserves in the province. The estimated 'census' count and the 'estimated net missed persons' in each reserve were then summed to create an 'estimated' population for the incompletely enumerated Indian reserves.

2.5 Estimates of census net undercoverage by single year of age and sex

2.5.1 Provinces and territories

Census coverage studies provide undercoverage and overcoverage estimates for large areas, such as provinces and territories and for large domain, such as age-sex combinations at the national level. However, the Demographic Estimates Program requires estimates of census net undercoverage by single year of age and sex for each province and territory. These are derived using various procedures which include:

1. direct survey to produce estimates of total census net undercoverage by province and territory (provincial and territorial marginal totals);
2. a calibrated curve smoothing method to produce the national estimates of census net undercoverage by age and sex (national age and sex marginal total);
3. Empirical Bayes regression model to create estimates of census net undercoverage at the provincial and territorial level, for broad age and sex categories;
4. synthetic model to generate estimates of census net undercoverage by single year of age;
5. raking ratio procedure to ensure the consistency of estimates of census net undercoverage by single year of age and sex with the provincial and territorial totals and national age and sex totals.

Marginal totals—provincial and territorial margin

The estimation procedures require two marginal totals: the provincial/territorial estimate of census net undercoverage and the national total of census net undercoverage by single year of age for each sex. The provincial and territorial estimates of census net undercoverage can be found in the official population release on census coverage studies. These are assumed to be correct and all domain estimates will be constrained to these totals.

Marginal total—national age-sex margin

The direct survey estimates of national census net undercoverage by age and sex require some smoothing. The direct estimates of census net undercoverage rates by single year of age (for each sex) from the coverage studies are used as the input data. However, unless the census net undercoverage rates between consecutive ages are smoothed, these estimates cannot be used as marginal totals. There are a variety of methods that can be used to smooth the direct estimates; in 2001 a cubic smoothing spline was introduced. This approach implicitly assumes that consecutive age groups have relatively similar census net undercoverage rates and change between ages follows a smooth function.

The model assumes that the true census net undercoverage rates—defined as the ratio of net missed persons over the total population—are described by a smooth function of age. This approach is the continuous version of the discrete Whittaker-Henderson Graduation Method used in the 1996 population estimates (Gambino and Dick 2000).

The RRC publishes the census net undercoverage rates by selected broad age groups: 0 to 19, 20 to 29, 30 to 44, 45 to 69, and 70 years and over. The smoothing spline will produce different estimates for these groups. A common procedure to ensure that small domain estimates are in agreement with higher level aggregated data is to calibrate the estimates to the totals for the broader age groups (Ramsay 2000).

Small domain estimation—broad age groups

Detailed estimates of census net undercoverage by single year of age and sex within each province/territory are handled in a two-step procedure. First, an Empirical Bayes regression model is used to derive broad age group and sex estimates of census net undercoverage at the provincial and territorial level, and then a synthetic estimate is created for the single years within the broad age groups.

The objective of modelling the small domain estimates of census net undercoverage is to produce a series of estimates with a smaller mean square error than the direct estimates. However, as opposed to the direct survey estimate which is design unbiased, the modelling approach will introduce a bias for each estimate. Thus modelling the small domain estimates of census net undercoverage involves a trade-off between reducing the variance of each estimate and the bias introduced through the modelling process. One approach to ensuring that the more reliable direct survey estimates are used is to introduce an Empirical Bayes model similar to Fay and Herriot (1979). This procedure creates an estimate of census net undercoverage that is a combination of a model estimate and the direct survey estimate weighted by their respective variances. Note that since the individual sampling variances are used in the estimation, a more precise direct estimate would contribute much more to the final Empirical Bayes estimate than a direct estimate with low precision. This ensures that the model does not dominate estimates that are already considered reliable.

The Empirical Bayes regression model starts with the direct survey estimates of adjustment factors (ratio of true population to census population) with variances for four broad age groups (0 to 19, 20 to 29, 30 to 44, and 45 years and over) and sex for each province and territory. This approach is best described as a two-stage regression model: a sample model that describes the basic relationship between the estimated adjustment factors and the true adjustment factors, and a regression model that describes how the true adjustment factors vary with a set of underlying variables.

The regression coefficients are estimated using a weighted least squares method which ensures that direct estimates from the large provinces with small standard errors are respected more than direct estimates from the small provinces with large standard errors. Since the sample variance estimates are highly unstable, they are first stabilized using a generalized variance function. The estimated variances are then used and assumed to be without error in the subsequent step (for details, refer to Dick 1995).

Small domain estimation—synthetic expansion

The Demographic Estimates Program requires estimates of census net undercoverage by single year of age and sex for each province/territory. The Empirical Bayes estimates, discussed above, can produce estimates for broad age groups but it cannot work effectively with more detailed age categories due to a lack of sample. Too many domains would produce an estimate of zero since no sample was in the domain. To meet the requirements of the Demographic Estimates Program, a synthetic estimate must be introduced.

Estimated adjustment factors for four broad age groups and sex in each province/territory are produced as a result of the Empirical Bayes model. A synthetic model is then used to expand the estimates from the broad age groups to single ages. This model essentially assumes that a constant census net undercoverage rate within the four broad age groups is to be maintained. The adjustment factors ($\hat{F}_{jk}$) from the Empirical Bayes model allow for an easy synthetic estimate of missed persons for any single age a, within the age group k for area j using the following equation:

Equation 2.9:

$$\hat{M}_{jk_a} = C_{jk_a} \times (\hat{F}_{jk} - 1)$$

where

$\hat{M}_{jk_a}$ = net number of persons missed in single year of age a in province/territory j and broad age group k;

C_{jk_a} = number of persons counted in the census for single year of age a in province/territory j and broad age group k;

$\hat{F}_{jk}$ = adjustment factor produced by Empirical Bayes model by broad age group k.

Consistency adjustment

The small domain estimation will not be consistent with the marginal totals discussed above. Hence a raking ratio adjustment is used on the small domain estimates to ensure consistency with both the provincial totals and the national age-sex totals.

This procedure organizes the estimates of missed persons into a matrix with the single year of age estimates as the row and the province/territory estimates as the columns. The fixed marginal totals are then used for the single ages at the national levels and for the provincial totals. The Empirical Bayes estimates of the adjustment factors are then used to generate the synthetic estimates of missed persons for each province/territory. These estimates are then alternatively adjusted so that they sum to the row and column totals. Convergence is usually reached in about three iterations (for details, refer to Dick 1995).

These estimates are then used as the small domain estimate of missed persons. By adding them to the census counts, the Demographic Estimates Program can create a base population for generating the population estimates.

2.5.2 Subprovincial areas

Base populations for census metropolitan areas and census divisions are obtained by applying the corresponding provincial/territorial census net undercoverage rates, available by age and sex. This synthetic estimate assumes that within a province/territory and for a single year of age there is a constant census net undercoverage rate. For example, in B.C. a 20 year old male would be assumed to be missed at the same rate across the entire province. Late enumeration and non-enumerated Indian reserves and settlements were adjusted by adding the provincial/territorial estimates to the appropriate geographic regions. All figures sum to provincial/territorial and national totals.

2.6 Estimates of census net undercoverage by marital status (and age and sex)

2.6.1 Provinces and territories

At this point, estimates of census net undercoverage are available by single year of age and sex for each province and territory. However, the Demographic Estimates Program requires estimates of provincial/territorial census net undercoverage by age, sex and marital status.

The estimates of census net undercoverage (CNU) by single year of age, sex and marital status (historical and legal[18]) for each province/territory were modeled by Social Survey Methods Division of Statistics Canada. The method used to estimate the legal and historical marital status census net undercoverage is a two step raking ratio procedure. The estimation procedures require two marginal totals to be fixed. One margin was the previously released estimate of net missed person by single year of age and sex for each province/territory. The other margin used was from the coverage studies showing the estimated net missed persons by age for marital status (historical and legal) for each province/territory. Some minor modifications to the direct survey estimates were necessary to ensure they were calibrated to the previously released estimates of persons over age 14.

The raking ratio procedure was used in a two step process. The first step was to use the five-year age groups and the second step was to use these results to produce the single year estimates. The choice was to create a design matrix using the census distribution of marital status, for each sex, within each five-year age grouping. The resulting distribution of marital status within an age group determined the structure of the design matrix. If any single marital status estimate in any five-year age group constituted at least 3 percent of the total for that age group, then the design matrix was set to one, otherwise it was set at zero. These conditions imply that the various categories are independent of each other if they meet this threshold. Finally, for each province/territory and sex, the legal marital status used the identical design matrix derived from the historical marital status. Convergence was usually achieved in about ten iterations. After convergence, the estimated CNU rates were checked, and if any estimate had a CNU rate exceeding 50 percent then the initial value for that cell was re-set to zero and the raking procedure re-run.

Single year of age estimates by province and territory were calculated by a second raking ratio procedure. Again two marginal totals had to be fixed. One margin was the counts of previously released single year of age by sex while the other margin was the historical and legal marital status estimates for each sex by province/territory. The initial counts were the five-year age estimates. The raking procedure usually converged in about ten iterations. The final single year estimates were then checked to ensure that for every single age the explicit constraint between legal and historical marital status was maintained. This constraint means that for the adjusted population, the difference between the historical marital status married/common law and legal marital status married is greater than or equal to zero. On the other hand, the difference between the historical status and legal status for the other category (divorced, separated, single and widowed) is less than or equal to zero. The estimates were then rounded to the nearest integer.

18. Legal marital status refers to the marital status of the person under the law. It includes these five categories: married (and not separated); widowed (including living common law); separated (including living common law); divorced (including living common law); and single (including living common law). All persons aged less than 15 are considered as single.

Chapter 3

Births and deaths

Births and deaths have been recorded on a regular basis since 1921. Since the registration of births and deaths is a legal requirement in Canada, these data are readily available and are generally of very high quality[19].

Births are added to the base population, and deaths, excluding stillbirths, are subtracted from the base population (along with the migration components) to generate the population estimates. The term "births" is commonly used to mean "live births", and the term "deaths" refers to "deaths excluding stillbirths."[20] This chapter contains information about the data sources for births and deaths and the methods of producing preliminary estimates.

3.1 Data sources and relevant concepts

Information on births and deaths is obtained from the vital statistics databases maintained by Statistics Canada's Health Statistics Division. The databases are created in collaboration with the provincial and territorial ministries responsible for recording vital statistics. Under provincial and territorial vital statistics acts (or equivalent legislation), the registration of all live births and all deaths within their jurisdictions is compulsory. The central vital statistics registry in each province and territory provides Statistics Canada with birth and death registration data. The data are available at the national, provincial/territorial, census division (CD) and census subdivision (CSD) levels.[21]

The universe of the birth database includes births to Canadians, landed immigrants and non-permanent residents in Canada. It also includes births to Canadians in some American states. The universe of the death database comprises deaths of Canadian residents, landed immigrants and non-permanent residents in Canada. Deaths of Canadian residents in some American states are also included. Births and deaths of foreign residents are excluded.

19. Undercoverage of births, due to late registration, births to Canadian residents outside Canada, non-registration of births, and deaths within days of birth, is thought to be minimal. The same is true for undercoverage of deaths, owing to late registration, delayed or missing registrations due to unidentified bodies, and deaths of Canadians outside Canada. Deaths of Canadians serving in the Armed Forces outside Canada are not included in Statistics Canada's databases because they are not registered by the provinces and territories. Overcoverage of births and deaths is not considered significant and is not measured.
20. For a detailed description of the birth database and the death database maintained by the Health Statistics Division, refer to the following links to Statistics Canada's official website: "Vital Statistics - Birth Database" and "Vital Statistics - Death Database."
21. Data for births and deaths from the vital statistics databases are not available for census metropolitan areas (CMAs). The Demography Division derives counts for CMAs using data for more detailed levels of geography.

3.2 Birth and death estimates, Canada, provinces and territories

3.2.1 Levels of estimates

There are three levels of estimates: preliminary, updated and final. The source of these series is the vital statistics data provided by the Health Statistics Division. For final estimates, the numbers of births and deaths are taken directly from vital statistics. Vital statistics are received annually, with a lag of two to three years relative to our reference period. For preliminary and updated estimates, we use the ratio method described below.

3.2.2 Final estimates

Because of the completeness of the vital statistics databases, the data provided by the Health Statistics Division (HSD) for demographic estimates requires very few adjustments. When both place of birth or death and place of residence are available in the data files, place of residence is preferred for reporting the event for the purposes of population estimates as well as vital statistics health indicators. When a woman gives birth outside her province (i.e., in a province other than her province of residence), the birth is counted in the mother's province of residence. Similarly, if a person dies outside his or her province of residence, the death is counted in the province of residence. This is because the birth or death affects the population of the province of residence and not that of the province where the event occurred.

Using the component method, the number of births that occurred between the date on which the base population was established and the reference date is simply added to the base population and the number of deaths is subtracted from the base population.

3.2.3 Preliminary estimates[22]

Since vital statistics are not available at the time the preliminary population estimates are prepared, births and deaths are estimated using the ratio method. The number of births is estimated with fertility rates by the mother's age group. The number of deaths is estimated with death rates by age group and sex. To produce preliminary or updated estimates, we use ratios based on the most recent year for which final estimates are available. This method is not used for Quebec and British Columbia, since their statistical agencies provide always their most recent estimates (see below for details). The preliminary estimates are produced on a quarterly basis.

The following formulas are used to generate quarterly preliminary estimates of births and deaths at the provincial/territorial level.

22. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.

For births:

By quarter for each province and territory, by mother's age group at the beginning of the quarter:

Equation 3.1:

$$B^{a_M}_{(t,t+2)} = \left(\frac{{}^{f}B^{a_M}}{{}^{f}P^{a_F}} \right) \times P^{a_F}_{t}$$

where

(t,t+2) = interval between the date of the previous reference period, time t and the reference date of the estimate, time t+2. In this case, the interval is a quarter;

$B^{a_M}_{(t,t+2)}$ = estimate of the number of births to mothers M aged a between time t and time t+2;

${}^{f}B^{a_M}$ = number of final births f to mothers M aged a during the corresponding quarter;

${}^{f}P^{a_F}$ = estimate of the population of females F aged a at the beginning of the corresponding quarter for which final birth estimates f are available;

$P^{a_F}_{t}$ = estimate of the population of females F aged a at the beginning of the quarter for which an estimate is required (time t);

a = age at the beginning of the quarter by five-year age group.

Then, the estimate of the total number of births is calculated as follows:

Equation 3.2:

$$B_{(t,t+2)} = \sum_{a} B^{a_M}_{(t,t+2)}$$

For deaths:

By quarter for each province and territory, by age at the beginning of the quarter, for each sex:

Equation 3.3:

$$D^{a,s}_{(t,t+2)} = \left(\frac{{}^{f}D^{a,s}}{{}^{f}P^{a,s}} \right) \times P^{a,s}_{t}$$

where

(t,t+2) = interval between the date of the previous reference period, time t and the reference date of the estimate, time t+2. In this case, the interval is a quarter;

$D^{a,s}_{(t,t+2)}$ = estimate of the number of deaths of persons aged a and sex s between time t and time t+2;

${}^{f}D^{a,s}$ = number of final deaths f of persons aged a and sex s during the corresponding quarter;

$^{f}P^{a,s}$ = estimate of the population of persons aged a and sex s at the beginning of the corresponding quarter for which final death estimates f are available;

$P_t^{a,s}$ = estimate of the population of persons aged a and sex s at the beginning of the quarter for which an estimate is required (time t).

a = age at the beginning of the quarter by five-year age group.

The estimate of total number of deaths is then calculated as follows:

Equation 3.4:

$$D_{(t,t+2)} = \sum_{s} \sum_{a} D^{a,s}_{(t,t+2)}$$

Monthly estimates are calculated using the most current final monthly distribution of births (or deaths).

The preliminary estimates cannot be finalized until the Health Statistics Division releases the vital statistics. Demography Division revises its estimates in the fall of each year.

3.2.4 Special treatment for preliminary estimates for Quebec and British Columbia

Quebec and British Columbia submit their most recent estimates of births and deaths to Statistics Canada. Those estimates are based on counts available from their respective vital statistics offices. Their figures are used to produce preliminary estimates for Statistics Canada's publication. For the final estimates, the two provinces birth and death figures are taken from the vital statistics compiled by the Health Statistics Division.

3.3 The use of birth and death estimates in other population estimates

Information about births and deaths from the Canadian vital statistics databases is used to produce other types of population estimates. For example, birth and death data (in addition to the other components of population change) are used to produce the population estimates by age and sex, which will be discussed in Chapter 9.

Chapter 4

Immigration

This chapter provides the information on the data sources regarding immigration, and the methods used to produce estimates of immigrants by age and sex, by province and territory. Information on the other four components of international migration can be found in the subsequent chapters.

4.1 Data sources and relevant concepts

The immigrant population refers to people who are landed immigrants in Canada. A landed immigrant is a person who is not a Canadian citizen by birth, but has been granted the right by immigration authorities to live in Canada on a permanent basis.

In Canada, immigration is regulated by the *Immigration and Refugee Protection Act* (IRPA) of 2002. This statute superseded the *Immigration Act*, which was passed in 1976 and amended more than 30 times in the years thereafter. Under the *IRPA*, there are three basic categories of permanent residents[23]: the economic class, the family class, and the protected persons category (or refugees).

Citizenship and Immigration Canada (CIC) collects and processes immigrants' administrative files. It then provides Statistics Canada with information from Field Operational Support System (FOSS) files. The information is used to estimate the number and characteristics of people granted permanent resident status by the federal government for a given period. Immigrants are usually counted on or after the date on which they are granted permanent resident status or the right to live in Canada. For Demography Division, the terms immigrant and permanent resident are equivalent.

23. Children born abroad to Canadian parents who are out of the country are, by definition, Canadian citizens, and therefore are not included in estimates of immigration. Included however, are those persons who change status from non-permanent residents (i.e., permit/authorization holders or refugee status claimants) to landed immigrant status from within Canada. Although their migrations do not involve crossing Canadian borders, they are counted as non-permanent residents, upon their initial entry to Canada.

4.2 Estimates of immigration, Canada, provinces and territories

Maintaining Canadian immigration statistics is statutory. Measuring the number of immigrants entering Canada in a given period is straightforward, and adjustments to the data are not required. Information is available for each person entering Canada under landed immigrant status from CIC's administrative file.

Each month, CIC makes available to Statistics Canada, a data file containing the records of landed immigrants for the previous month, as well as any additions or updates to data already received. Given that there are typically few changes to the CIC data, the differences between preliminary and final estimates are very small.

For provincial and territorial level estimates, the file obtained from CIC identifies the province or territory of intended destination upon arrival, rather than the province or territory in which the immigrant actually settles. In a small number of cases, information on the province of destination is lacking. For these cases, the province of destination is distributed proportionately between the provinces and territories according to the distribution observed from immigrants for whom the information is available.

4.2.1 Immigration estimates by age and sex

The distribution of immigrants by age and sex is also straightforward, as these variables are available from the CIC file. The distribution only requires basic tabulation by age and sex. In the event of missing information, these cases are prorated according to the distribution for immigrants for whom the information is available.

4.2.2 Levels of estimates

The difference between preliminary[24] and final postcensal estimates lies in the timeliness of the source used to estimate this component. Since the FOSS file is continually being updated, new calculations are carried out each year to update the immigration estimates. Immigration estimates are preliminary in the first year and finalized the following year.

24. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.

Chapter 5

Net non-permanent residents

Non-permanent residents (NPRs) are persons who have been legally granted the right to live in Canada on a temporary basis under the authority of a temporary resident permit, along with members of their family living with them. This chapter provides the information on the data sources on NPRs, and the methods used to produce estimates of NPRs by age, sex, province and territory.

5.1 Data sources and relevant concepts

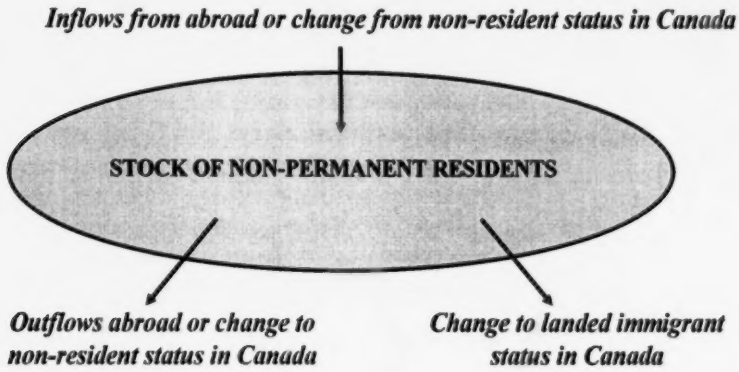
The data required to produce the NPR estimates are obtained from Citizenship and Immigration Canada's (CIC) Field Operations Support System (FOSS) files. They include data on visitor permits, work permits, student permits, special temporary residents' permits, refugee status claims, landings[25], deportations, applications for landing and proof of identity information for persons making refugee status claims. The information is used to estimate the number and characteristics of people granted non-permanent resident status by the federal government.

Non-permanent residents include foreign workers, foreign students, the humanitarian population and other temporary residents. The humanitarian population includes refugee claimants and temporary residents who are allowed to remain in Canada on humanitarian grounds and are not categorized as either foreign workers or foreign students. For Demography Division, the terms non-permanent resident and temporary resident are equivalent. Children born in Canada to parents of non-permanent resident status are considered as Canadians by birth and have all rights and privileges associated with citizenship.

Movements into and out of the NPR population are referred to as "flows", as are status changes from permit holders to refugee status claimants or from permit holders or refugee status claimants to landed immigrant status[26]. "Stocks" represent the number of NPRs residing in Canada at any given time. As illustrated in Figure 5.1, from the most general perspective, persons can enter into the current stock of NPRs (i.e., "inflows") either from abroad or from non-resident status within Canada (e.g. visitors). When a person leaves the NPR population (i.e., "outflows"), it is assumed that they have either left the country, become a non-resident, or become a landed immigrant.

25. "Landings" refers to lawful permission to establish permanent residence in Canada.
26. Persons leaving the NPR population by obtaining landed immigrant status change the size of the NPR population, but have no impact on the size of the total population of Canada, as these persons become a part of the permanent resident population.

Figure 5.1
Non-permanent residents flows by broad type


Inflows from abroad or change from non-resident status in Canada
STOCK OF NON-PERMANENT RESIDENTS
Outflows abroad or change to non-resident status in Canada
Change to landed immigrant status in Canada

In the Demographic Estimates Program, NPR stocks are included in the total population estimates for a given date. Meanwhile "net change" in NPR stocks over a certain period of time is included as a component of demographic growth. "Net change" in NPR represents the variation in the number of NPRs between the two dates.

NPRs may be in Canada under any one of the following situations:

(i) as holders of only permits (they may concurrently hold more than one type);
(ii) as persons who have only ever lodged refugee status claims; and
(iii) as refugee status claimants who also possess one or more of the types of permits that would qualify them for NPR status.

In cases where refugee status claimants concurrently hold some type of permit (e.g., a refugee status claimant is granted an employment permit to help support themselves and their dependants within Canada), the refugee status supersedes the other NPR status. Consequently, references to "permit holders" excludes persons who have also made refugee status claims. The two major NPR sub-groups, permit holders and refugee status claimants, are used for estimation purposes. Dependants who were born abroad to members of these two sub-groups of NPRs are also included; however children born in Canada to NPRs are not, this population being covered by the birth component (see Chapter 3).

All persons in possession of a permit or claiming refugee status are assigned a Client Identification Number (CID) by Citizenship and Immigration Canada (CIC). This identification number is a key variable as it is unique for each person holding a permit or claiming refugee status and is maintained for every document issued by CIC. It is used in the production of estimates of NPRs.

Anyone who received non-permanent resident status prior to the reference date is counted in the NPR population. For refugee claimants, the date of their application is used as the date they receive NPR status. For permit holders the effective date is typically the start date of their permit. Permit holders and refugee claimants are excluded from the population if their permit has expired, if they receive permanent resident status, or if they are deported. In addition, refugee claimants are excluded if their file has been inactive for two years.

5.2 Estimates of stocks of non-permanent residents

The two major subgroups of the NPR population (permit holders and refugee status claimants) are administratively different; therefore, their estimates must be produced separately. The methods used in the production of estimates for permit holders are discussed first, followed by those used for refugee status claimants.

5.2.1 Stocks of permit holders

Permit holders (PHs) must either have had an official document signed in advance or have it signed upon entry to Canada, allowing them to reside in Canada on a temporary basis. Therefore, a person is considered part of this population if they possess a valid employment, student or Ministerial document on the reference date for the estimates. This means that the document must have been in effect prior to the reference date and valid until or past the reference date. If a person possesses more than one valid permit, the information from the permit for which the effective date[27] is closest to the reference date is used.

Since dependants of permit holders are not required to obtain their own permits (though some do), the number of permit holders and their dependants is equal to the total number of persons covered by the documents.

Status as a permit holder terminates upon the expiry of the valid document, or when a PH is granted landing status (i.e., becoming a permanent resident), deported, or claims refugee status.[28] Though the latter case affects the populations of the sub-groups of NPRs, the total NPR population is unaffected by this type of change in status.

The province of residence for each PH is obtained from the valid document. If a person has more than one valid permit, then the province of residence is taken from the permit for which the effective date is closest to the reference date. In some circumstances, the

27. The effective date is the date (or an approximation of) when the PH entered Canada as an NPR, or the date as of which a permit is extended.
28. Occasions arise where there appears to be a short interruption in a PH's temporary stay in Canada. That is to say, a permit expires and there is a brief time lag between the expiry date and the effective date of the next permit. Since this is likely due to administrative delays in the issuance of permits and extensions, interruptions of less than 31 days are disregarded and these persons are considered as having been continuously residing in Canada.

province of residence is not indicated on the document. In these cases, the province of the Canada Immigration Centre where the permit was issued is used. There are a small number of cases for which there is neither a province of residence nor a Canada Immigration Centre indicated. These cases are prorated according to the provincial distribution of PHs for which the province of residence is known or has been derived. In all cases, the province of residence for the principal PH is assumed for the dependants.

For each province/territory the stocks of PHs is a simple accounting of all valid permit holders and their dependants and can be expressed as follows:

PH = sum of all PHs and their dependants covered by the valid document on the reference date, as defined above

5.2.2 Stocks of refugee status claimants

A person is considered part of the refugee status claimant (RSC) population once a claim has been filed at a Canada Immigration Centre. Effective with the proclamation of Bill C-86 on February 1, 1993, each person claiming refugee status is treated as a separate "case" (i.e., all claimants, including dependants are considered of separate identity and each is represented by a separate case, identified by a unique CID). Under the previous legislation (Bill C-55), it was not necessary for dependants to submit independent claims; therefore the number of RSCs was equal to the total number of persons represented by the claim of principal claimants.

Those persons who held legal temporary status in Canada as PHs prior to making a refugee status claim are included in the PH population until the date of their refugee status claim, at which time they are considered to be RSCs.

Since refugee status claims do not have an expiry date, a withdrawal of a claim, a deportation or being granted landing status has to occur for a claimant to leave the NPR population. For this reason, some assumptions are necessary to ascertain if a claim is still active.

For those RSCs having applied for landed immigrant status:

(i) it is assumed that they will leave the NPR population 2 years after their application for landing date unless they do not possess acceptable proof of identity (see (ii) below). This assumption is based on the fact that these applications typically take no longer than 2 years to process;

(ii) it is assumed that if they do not have valid proof of identity[29], they will leave the NPR population 5 years after the date of application for landing. This

29. Cases of persons arriving in Canada without proof of identity may arise in situations where they have left a country involved in civil unrest or war, with little or no belongings. Depending on the circumstances in their last country of residence, they may not be capable of obtaining appropriate identification even after they have entered Canada.

assumption is in accordance with the recommendation of the Minister of Immigration.

For those RSCs who have not applied for landed immigrant status:

(iii) it is assumed that they will leave the NPR population 2 years after their last communication with CIC. Records of every communication with each client are kept, whether it be a request for information on their status, to file a new application, to make an administrative change (e.g., change of address), etc. If there has been no activity over a 2 year period for an RSC, it is assumed that they are no longer an NPR. The assumption is that they either have left the country, or are deceased.

The province of residence for each RSC may be obtained from the claim record. If it is missing or invalid, if it exists, then information from permit is used as a proxy (see previous section). If there is no information on the province of residence from these sources, then as with PHs, the province of the Canada Immigration Centre where the claim was made is used. If this is not available, these remaining cases are prorated according to the provincial distribution for RSCs where the province of residence is available (or derived).

Given the above, stocks of RSCs can be expressed as follows:

For each province/territory:

(a) for claims submitted under Bill C-55 (between January 1, 1989 and January 31, 1993), the total number of persons covered by the active claim on the reference date; plus

(b) for those submitted under Bill C-86 (effective February 1, 1993), the number of active individual claims on the reference date.

5.2.3 Level of estimates

The difference between preliminary[30] and final estimates lies in the timeliness of the source used to estimate this component. Estimates of the net number of NPRs are revised on an annual basis. Non-permanent resident (NPR) estimates are preliminary the first year and updated the following year. They become final two to three years after the reference year.

30. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.

5.3 NPR estimates by age and sex

5.3.1 Permit holders

Though permits include information on the number of persons the document covers, data related to the age and sex are only available for the principal PH. Therefore, the information for principal applicants is obtained directly from the CIC files, while estimation is required for characteristics of their dependants.

Since data are available for all RSCs who filed under Bill C-86 (i.e., one application equals one person), the age and sex distributions of dependants of principal claimants who also hold permits, are used for the dependants of PHs.

5.3.2 Refugee status claimants

Since data are available for all RSCs who filed under Bill C-86, the data on sex and age for these claimants are obtained directly from their valid claim.

5.4 Net non-permanent residents as a component of change of total population estimates

The net number of NPRs for a given period is calculated by subtracting the stock of NPRs at the beginning of the period from the stock at the end of the period.

Equation 5.1:

$$\Delta NPR_{(t,t+i)} = NPR_{(t+i)} - NPR_t$$

where

$(t,t+i)$	=	interval between the date of the previous reference period, time t, and the reference date of the estimate, time t+i;
$\Delta NPR_{(t,t+i)}$	=	change in NPR stocks between time t and time t+i;
$NPR_{(t+i)}$	=	NPR stocks at time t+i;
NPR_t	=	NPR stocks at time t.

In the Demographic Estimates Program, net change in NPRs is one of the components added to the base population to produce population estimates at a given date.

Chapter 6

Emigrants, net temporary emigrants and returning emigrants

Emigration refers to citizens or immigrants who leave the country to establish a residence in another country. This residence abroad may be intended as permanent or temporary. Emigration results in a loss of population. Taking into account all aspects of Canadian emigration, the calculation of population involves the estimation of emigrants, net temporary emigrants and returning emigrants.

Unlike immigration, there is no legal provision in Canada to maintain records for persons leaving the country either on a temporary or permanent basis. Therefore, estimates of the number of emigrants and persons living temporarily abroad and their characteristics must be derived through secondary sources such as Canadian administrative files or immigration statistics of the United States.

As emigration components are the most difficult components to estimate, the methods of estimation are constantly evolving in an attempt to produce more accurate emigration estimates, given available sources of information and feasible methods. This chapter presents current methods used to produce the estimates of the three components of population change resulting from emigration.

6.1 Data sources and relevant concepts

Despite recommendations by the United Nations to establish a universal definition of international migrants, as well as a uniform method of recording information regarding emigration and immigration, not much progress has been made (United Nations, 1998). This has consequences on the ability to compare migration statistics between different countries, as well as the potential to use other nations' international migration statistics.

For the purpose of estimating Canada's population, the following three components relating to emigration are described. Emigration estimates require a distinction between those persons establishing a permanent residence in another country (i.e., emigrants), those persons living temporarily abroad (i.e., net temporary emigrants), and finally the portion of emigrants who have returned to Canada (i.e., returning emigrants). Different data sources and methods are necessary for the two types of emigrants, as well as for returning emigrants. Estimates of emigrants, net temporary emigrants and returning emigrants are provided as separate components in publications on Canadian population estimates.

Emigrants are Canadian citizens or immigrants who have left Canada to establish a permanent residence in another country (sometimes referred to as "permanent emigration"). In the census, permanent emigration involves severing residential ties with

Canada and acquiring permanent residency in another country. For example, persons traveling to the United States may be considered permanent emigrants if they acquire permanent resident status there, but are considered temporary emigrants if they still hold a visa or are on visitor status. Permanent emigration is a misnomer as it is not necessarily irreversible; emigrants can always decide to return to Canada.

Emigrants are estimated from administrative sources in terms of the "gross flow" of migrants out of Canada. The Office of Immigration Statistics, U.S. Department of Homeland Security provides data on Canadians who acquire landed immigrant status in the U.S. This data source is used in estimating emigration to the United States. In order to estimate emigration to other countries, information on notification of departure from the Canada Child Tax Benefit (CCTB) program and tax data from Canada Revenue Agency (CRA) is used.

Canada Child Tax Benefit data (CCTB), combined with personal income tax data, cover emigration better than tax data alone. The CCTB program identifies emigrants through their tax return and/or from their notification of departure to CRA. The notification of departure can cover, in part, families that will not necessarily file a tax return for the year of their departure. The combination of CCTB and tax data allows for an estimate of child and adult emigrants and, using Demography Division's population estimates, provides an estimate of the non-covered CCTB child population. Tax data on filers and their dependants are provided by Income Statistics Division (ISD) of Statistics Canada, based on files received from CRA.

Some people leave Canada to live temporarily in another country while not maintaining a usual place of residence in Canada. Others who were temporarily outside Canada return. The net result of those departures and returns is the component known as "**net temporary emigration**". Data from the Reverse Record Check (RRC), the most important census coverage study, are used to estimate the number of persons leaving the country temporarily; while data from the census, combined with Demography Division's estimates of returning emigrant, are used to estimate the number of temporary emigrants returning.

Temporary migration constitutes many movements in and out of the country. This would not concern the overall population estimates if the net effect was nil or negligible. However, a census coverage study (Reverse Record Check) in 1996 has shown that this is not the case. The decision to account for persons living temporarily abroad was based on evaluations of the 1996 error of closure for postcensal population estimates and its components against Reverse Record Check estimates. These evaluations concluded that omitting departures of Canadians for temporary residence abroad and their consecutive returns to Canada has an important impact on the quality of the estimate of the country's population (Michalowski, 1999). Including net change in temporary emigration in the Demographic Estimates Program is a practice that started in 1998, with revisions to estimates back to 1996.

Returning emigrants are Canadian citizens or immigrants having previously emigrated from Canada and subsequently returned to Canada to re-establish a permanent residence. Again, data from the CCTB program and from CRA's T1FF[31] are used in estimating returning emigrants.

6.2 Estimates of emigrants, net temporary emigrants and returning emigrants

6.2.1 Emigrants

The number of emigrants is estimated using data from the Office of Immigration Statistics, U.S. Department of Homeland Security, data collected by the Canada Child Tax Benefit (CCTB) program, and data from the T1 Family File (T1FF). The first source is used to estimate emigration to the United States while CCTB data are used to estimate emigration to other countries. The estimates of the number of child emigrants have to be adjusted because the CCTB program is not universal and does not provide direct information on the number of adult emigrants. As a result, four adjustment factors are used to take into account:

1. the incomplete coverage due to a delay in the receipt and processing of the files of children eligible for the CCTB. Emigration estimates are usually finalized two years after the reference period. Based on historical files, it seems that CCTB data recordings reached a significant level two years after the year of emigration and that it takes four years for CCTB files to become complete; thus an adjustment is made if the estimates are finalized after two years;

2. the program's partial coverage, that is, people who do not apply for the CCTB or are not eligible[32];

3. the differential propensity to emigrate between children who are eligible for the CCTB and children who are not; and

4. the differential propensity to emigrate between adults and children.

Each of these adjustment factors makes use of CCTB and tax data from CRA, and the estimation methods for child and adult emigrants are described in the next section.

31. The T1 Family File (T1FF) is derived from the Canada Revenue Agency (CRA) T1 file by Income Statistics Division of Statistics Canada.
32. Canadian citizens, permanent residents and non-permanent residents who have been in Canada for a year and subject to Canadian taxation are eligible for the CCTB.

Emigrant children

The CCTB file provides numbers of dependant children (under 18 years of age) whose parents, who are eligible and have applied for the CCTB, have become non-residents as defined by CRA. These data are available on a monthly basis but provided by CRA to Statistics Canada annually for each province and territory. As mentioned, the CCTB program is not universal; therefore the data are incomplete in terms of estimation of total child emigrants and require adjustments.

The first correction factor is to take into account the incompleteness of the emigration information from the CRA's CCTB data files due to delays in the recordings of an emigration. Based on comparisons done with files that were two, three and four year after the reference period, it appeared that the CRA files could be considered complete about four years after the reference period. Therefore, an adjustment is made to the data when they are finalized, which is only two years after the reference period. The same adjustment factor is applied to the monthly data of each province/territory.

The second correction factor, that is the adjustment for partial coverage, corrects for the absence of universality of the program. The adjustment starts with a correction factor that is applied to the population registered for the CCTB program to account for its shortfall in total coverage: the ratio of the number of children registered for the CCTB program to the number of children in the total population, as estimated by Demography Division for each month and province and territory.

The third correction factor is used to adjust the differential "emigration propensity" between children registered for the CCTB program and that of all children for each province and territory, on an annual basis. This factor is obtained by comparing the emigration rates of CCTB-eligible children with the rates for all children (aged 0 to 17). This factor is calculated for each province and territory and is based on the last three available years of the T1FF. To eliminate variations due to small numbers in each of the Atlantic provinces, the differential propensity factor is estimated for these provinces as a group. In addition, the estimated factor for Canada is used for the territories.

The formula to estimate child emigrants by province/territory is as follows:

For each province/territory:

Equation 6.1:

$$ {}^{CCTB}_{j}EM^{0-17} = {}_{j}CCTB\_EM^{0-17} \times \frac{1}{{}^{CCTB}_{j}\Re} \times {}_{j}G \times D $$

where

${}^{CCTB}_{j}EM^{0-17}$	=	emigrant children aged 0 to 17 from province/territory j based on CCTB data;
${}_{j}CCTB\_EM^{0-17}$	=	number of CCTB/Tax children moving from province/territory j;
${}^{CCTB}_{j}\Re$	=	coverage rate of CCTB program for province/territory j;
${}_{j}G$	=	adjustment factor for emigration propensity of non CCTB/Tax children for all provinces except the Atlantic provinces where the factor is measured as a whole and in the territories where the Canadian factor is used;
CCTB	=	CCTB data;
D	=	adjustment factor for the incompleteness of emigration data from CCTB files due to delays in the recordings of an emigration.

The equation above includes the adjustments for incomplete coverage of CCTB-registered children (Equation 6.2), for the differences in the propensities of CCTB-registered and non-registered subpopulations to emigrate (Equation 6.3) and for the incompleteness of emigration data due to delays in the recordings of an emigration (Equations 6.4a to 6.4c).

The coverage rates are calculated on a monthly basis as follows:

For each province and territory:

Equation 6.2:

$$ {}^{CCTB}_{j}\Re = \left[\frac{{}^{CCTB}_{j}P^{0-17}}{{}^{Dem}_{j}P^{0-17}} \right] $$

where

${}^{CCTB}_{j}\Re$	=	coverage rate of CCTB program for each province/territory j;
${}^{CCTB}_{j}P^{0-17}$	=	number of children aged 0 to 17 years registered for the CCTB program in each province/territory j;
${}^{Dem}_{j}P^{0-17}$	=	population estimated by Demography Division of children aged 0 to 17 years in each province/territory j.

The differential propensity to emigrate is obtained by dividing the emigration rates for all children by the emigration rates for children registered for the CCTB program, as follows:

For each province and territory:

Equation 6.3:

$$ {}_jG = \frac{{}_j\text{Rate\_Tax\_EM}^{0-17}}{{}_j\text{Rate\_CCTB\_Tax\_EM}^{0-17}} $$

$$ {}_jG = \frac{\left[\frac{{}^{\text{Tax}}_j\text{EM}^{0-17}}{{}^{\text{Tax}}_j\text{P}^{0-17}}\right]}{\left[\frac{{}^{\text{CCTB\_Tax}}_j\text{EM}^{0-17}}{{}^{\text{CCTB\_Tax}}_j\text{P}^{0-17}}\right]} $$

where

${}_jG$	=	adjustment factor for emigration propensity of non-CCTB/Tax children for each province[33] j of origin;
${}_j\text{Rate\_Tax\_EM}^{0-17}$	=	emigration rate of all children from each province[33] j;
${}_j\text{Rate\_CCTB\_Tax\_EM}^{0-17}$	=	emigration rate of CCTB-registered children from each province[33] j, derived from the T1FF by ISD;
${}^{\text{Tax}}_j\text{EM}^{0-17}$	=	emigration of all children from province, region or Canada j, according to income tax data;
${}^{\text{Tax}}_j\text{P}^{0-17}$	=	ISD estimate of children in province, region or Canada j based on income tax files;
${}^{\text{CCTB\_Tax}}_j\text{EM}^{0-17}$	=	emigration of CCTB-registered children from province, region or Canada j, derived from the T1FF by ISD;
${}^{\text{CCTB\_Tax}}_j\text{P}^{0-17}$	=	CCTB-registered children in the population of the income tax file provided by ISD by province, region or Canada j.

We use delay factors to compensate for the incompleteness of the CCTB data file on child emigration due to registration delays. Based on past trends, we have estimated that it takes about four years for the CCTB data file to be near complete. Since we usually finalize the estimates two or three years after the reference year, we use a delay factor to project what the data would be one or two years ahead.

Table 6.1 illustrates the calculation of the delay factors; all numbers in the table are fictitious. In this table, we suppose that the production year is 2011, therefore the t-3, t-2 emigration year in the first column corresponds to 2008/2009 and the t-4, t-3 emigration year corresponds to 2007/2008, etc. The subsequent three columns of the table show

33. Atlantic provinces all have the same factor/rate and the Canadian factor/rate is used for the territories.

numbers of child emigrants by year of emigration and by duration between the emigration year and the data file year. It can been seen from Table 6.1 that numbers of child emigrants for the emigration year t-4, t-3 from the data file that is two years after the emigration year is 7,900 (duration of two years between the emigration year and the data file year) and the number is 8,500 from the data file that is three years after the emigration year (duration of three years between the emigration year and the data file year). Our objective is to estimate what the t-4, t-3 and also the t-3, t-2 emigration year numbers would be in a file that is four years after the emigration year.

Table 6.1
Calculation of the delay factors: An illustration for production year 2011

Year of emigration	Number of child emigrants in the CCTB file after			$D_{(2,3)}$	$D_{(3,4)}$	$D_{(2,4)}$
	2 years	3 years	4 years			
t-7, t-6 (2004/2005)	6,300	6,750		1.071	1.052	1.138
t-6, t-5 (2005/2006)	6,700			1.082	1.048	1.142
t-5, t-4 (2006/2007)			8,200	1.090	1.045	1.124
t-4, t-3 (2007/2008)	7,900	8,500	Z=8,908	1.076		
t-3, t-2 (2008/2009)	7,700		Y=8,740			
Average factors (3 years)				1.083	1.048	1.135

Note: The data in diagonal are from the same data file year, as indicated by the different shades of grey. Darkest grey represents data from the 2008 file while lightest grey relates to 2011 data.

We first calculate delay factors for the previous emigration years. The delay factors (D) in the right-hand columns are derived from the numbers on the left. $D_{(2,3)}$ is the ratio between duration 3 and duration 2; while $D_{(3,4)}$ is the ratio between duration 4 and duration 3. For example, $D_{(3,4)}$ of 1.045 for the emigration year t-5, t-4 comes from the ratio between 8,200 and 7,850. This factor could be used to project one year ahead from duration 3 to duration 4. But, to project two years ahead from duration 2 to duration 4, equal to 1.124, we calculate the factor D(2,4) (the last column in Table 6.1) by multiplying the two most recent factors, D(2,3) equal to 1.076 and D(3,4) equal to 1.045.

The formulas are as follows:

Equation 6.4a:

$$D_{(2,3)} = \frac{CCTB\_EM^{t}_{(t-4,t-3)}}{CCTB\_EM^{(t-1)}_{(t-4,t-3)}}$$

Equation 6.4b:

$$D_{(3,4)} = \frac{CCTB\_EM^{t}_{(t-5,t-4)}}{CCTB\_EM^{(t-1)}_{(t-5,t-4)}}$$

where

$CCTB\_EM^{t}_{(t-x-1,t-x)}$ = the number of child emigration for the emigration year (t-x-1, t-x) from the CCTB file of year t.

If we want to project numbers for an emigration year that is three years before the production year, then we use the delay factor in the recordings of data on emigration from duration 3 to duration 4 (i.e., $D_{(3,4)}$).

If we want to project numbers for an emigration year that is two years before production year, then we use:

Equation 6.4c:

$$D = D_{(2,4)} = D_{(2,3)} \times D_{(3,4)}$$

In production, we use a three-year average of each of the $D_{(3,4)}$ and the $D_{(2,4)}$ factors based on the most recent factors available. For example, in year t, we will project the emigration number for the year t-3, t-2 (7,700) with a three-year average of the most recent $D_{(2,4)}$. The average is based on the factors: 1.138, 1.142 and 1.124. This average factor (1.135) will then be used to multiply the number of emigrants in year t-3, t-2 at duration 2 (7,700) and the results (Y=8,740) will be the number expected in the file that is four years after the emigration year. The same average calculation of the $D_{(3,4)}$ factor (1.048) is used to multiply the emigration number of year t-4, t-3 at duration 3 (8,500) to obtain the expected number (Z=8,908) in the file that is four years after the emigration year.

Data on Canadian adults and children who became immigrants in the United States are available from the Department of Homeland Security on an annual basis. The Homeland Security data provide quarterly flow of immigrants to the United States, including those emigrating from Canada. Since the Homeland Security data are of better quality than the estimates of emigration from the CCTB model, the CCTB estimates are replaced by the Homeland Security data for emigration to the United States. Combining the estimates of

child emigrants from the CCTB data and the Homeland Security data on children who moved to the United States, estimates of children emigrating to countries other than the United States can be derived as follows:

Equation 6.5:

$$^{OC}EM^{0-17} = {}^{CCTB}EM^{0-17} - {}^{HS}EM^{0-17}$$

where

$^{OC}EM^{0-17}$ = emigrant children to countries other than the United States;

$^{CCTB}EM^{0-17}$ = emigrant children based on CCTB data;

$^{HS}EM^{0-17}$ = emigrant children to the United States based on Homeland Security data.

Emigrant adults

Estimates for adults emigrating to the United States are taken directly from the Homeland Security data. As the CCTB program does not provide direct information on emigrant adults, hence an adjustment factor was used to estimate the number of adults emigrating to countries other than the United States, based on the emigration rate of children to other countries.

The formula for estimating the number of adult emigrants to a country other than the United States is:

Equation 6.6:

$$^{OC}EM^{18+} = \frac{^{OC}EM^{0-17}}{^{Dem}P^{0-17}} \times {}^{Dem}P^{18+} \times \sum_{t}^{t,t-2} \frac{RatioAC_t}{3}$$

where

$^{OC}EM^{18+}$ = emigrant adults to countries other than the United States;

$\frac{^{OC}EM^{0-17}}{^{Dem}P^{0-17}}$ = emigration rate of children moving to countries other than the United States;

$^{Dem}P^{18+}$ = population estimate of adults at the beginning of the month;

$RatioAC_t$ = ratio of adult emigration rate over child emigration rate at the reference date, time t;

t = the year index for the current reference date up to the two previous reference dates (t, t-1, t-2).

The RatioAC of adult migration rate over child migration rate is an average of ratios from two sources: tax and Homeland Security[34]. It is calculated for any given year as follows:

Equation 6.7:

$$\text{RatioAC} = \frac{\left({}^{\text{Tax}}\text{RatioAC} + {}^{\text{HS}}\text{RatioAC}\right)}{2}$$

where

${}^{\text{Tax}}\text{RatioAC}$ is the adult emigration rate divided by the child emigration rate taken from tax data. It is given by:

Equation 6.8:

$${}^{\text{Tax}}\text{RatioAC} = \frac{\left[\frac{{}^{\text{Tax}}\text{EM}^{18+}}{{}^{\text{Dem}}\text{P}^{18+}}\right]}{\left[\frac{{}^{\text{Tax}}\text{EM}^{0-17}}{{}^{\text{Dem}}\text{P}^{0-17}}\right]}$$

where

${}^{\text{Tax}}\text{EM}^{18+}$ = emigrant adults measured from tax data;

${}^{\text{Dem}}\text{P}^{18+}$ = estimate of adult population estimated by Demography Division (based on the average of start-of-period and end-of-period populations);

${}^{\text{Tax}}\text{EM}^{0-17}$ = emigration of all children according to income tax data;

${}^{\text{Dem}}\text{P}^{0-17}$ = estimate of children population estimated by Demography Division (based on the average of start-of-period and end-of-period populations);

and

${}^{\text{HS}}\text{RatioAC}$ is the adult emigration rate divided by the child emigration rate taken from the Homeland Security data. It is given by:

Equation 6.9:

$${}^{\text{HS}}\text{RatioAC} = \frac{\left[\frac{{}^{\text{HS}}\text{EM}^{18+}}{{}^{\text{Dem}}\text{P}^{18+}}\right]}{\left[\frac{{}^{\text{HS}}\text{EM}^{0-17}}{{}^{\text{Dem}}\text{P}^{0-17}}\right]}$$

where

34. Because of past legislative changes that have occurred in the US and other unforeseen events, sometimes the information was not available at production time. If either sources of the ratio of child emigration over adult emigration is missing or late for the production cycle of the estimates, then the past calculation of this ratio is used as a proxy.

$^{HS}EM^{18+}$ = emigrant adults measured from Homeland Security data;

$^{Dem}p^{18+}$ = estimate of adult population estimated by Demography Division (based on the average of start-of-period and end-of-period populations);

$^{HS}EM^{0-17}$ = emigration of all children according to Homeland Security data;

$^{Dem}p^{0-17}$ = estimate of children population estimated by Demography Division (based on the average of start-of-period and end-of-period populations).

Finally, the total number of adult emigrants equals:

Equation 6.10: $$EM^{18+} = {}^{OC}EM^{18+} + {}^{HS}EM^{18+}$$

where

EM^{18+} = *emigrant adults;*

$^{OC}EM^{18+}$ = *number of adult emigrants to countries other than the United States;*

$^{HS}EM^{18+}$ = *number of adult emigrants to the United States, according to Homeland Security data.*

The process of combining CCTB and Homeland Security data produces slightly inconsistent data in that the definitions of emigration used by the two data sources are not entirely the same. The Homeland Security definition of an immigrant to the United States (i.e., emigrants from other countries) is dependant on a person acquiring legal permanent resident status in their country. On the other hand, the CCTB usage does not subscribe to a definition of permanent resident status in another country to be considered an emigrant: one simply loses their resident status in Canada for income tax purposes. It is assumed that the difference between the two definitions is not sufficiently large to produce significant bias in the estimates.

Since the Homeland Security data do not provide information about the province of origin of Canadian emigrants, provincial/territorial level estimates of adult emigration must be derived from an alternate source. To estimate the adult emigrant distribution by province/territory of origin, the provincial/territorial distribution for child emigrants is assumed to be the same for the entire population.

Estimates of emigrants by age and sex

The choice of data source for each distribution (age, sex) was based on comparisons between the various data sources for emigration statistics (annual ISD data and Homeland

Security data, as well as the Reverse Record Check (RRC), which provides information on emigration every five years).

The national estimate of emigrants is first apportioned by sex. For emigrants to the United States, the breakdown is directly available from the Homeland Security data. For emigrants to other countries, the Homeland Security distribution is used as a proxy. This distribution is closer to what was measured with the RRC than is the ISD's estimates by sex.

Estimates of male and female emigrants are then apportioned by age groups, using ISD's distribution for the following five broad age groups: 0 to 17, 18 to 24, 25 to 44, 45 to 64, and 65 years and over. The Homeland Security data are used to further distribute the data into five-year age groups. Finally, Sprague's multipliers are applied to split the estimates into single years of age[35]. Since provincial/territorial level data are either unreliable or not available from these sources, the national age/sex distribution is used for all provinces and territories.

Levels of estimates

The difference between preliminary[36] and final estimates lies in the timeliness of the sources used to estimate this component. The same estimation method is used.

6.2.2 Net temporary emigration

Some people leave Canada to live temporarily in another country; others who were temporarily outside Canada return. The net result of those departures and returns is the component known as "net temporary emigration". The following steps are used to estimate their monthly numbers by province/territory:

1. The estimate of the number of departures at Canada level is derived from the Reverse Record Check (RRC), the most important census coverage study. The RRC provides an estimate of the number of people who left Canada temporarily[37] during an intercensal period and who are still out of the country at the end of the period.
2. The estimate of the number of temporary emigrants returning is done in two steps:
    a. The number of all returning emigrants for Canada is taken from the census. The census provides the number of persons who resided outside Canada at the previous census and who have since returned to the country during the intercensal period;

35. Sprague's multipliers are interpolation coefficients used to subdivide data. For a detailed description of the method of Sprague's multipliers, see Shryock, Siegel et al. 1976.
36. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.
37. Emigration is defined as temporary based on the intention of returning and the time spent outside the country.

b. From the estimate of all returning emigrants for Canada (census estimates Step 2a) we subtract Demography Division's estimate of returning emigrants. The resultant estimate is the number of temporary emigrants returning.

3. The estimate of net temporary emigration for Canada is derived by subtracting returning persons (Step 2b) from departures (Step 1).
4. The estimate derived in Step 3 is then distributed by province/territory according to the provincial/territorial distribution of departures of temporary emigrants based on the RRC[38]. The number for the Atlantic provinces is estimated as a group and redistributed proportionately to each province according to their respective population size. The same is done for the three territories.
5. The provincial/territorial estimates are disaggregated equally into annual estimates for each of the five years of the intercensal period. The monthly estimates are assumed to have a seasonal distribution. This distribution is modeled after the seasonal patterns observed for permanent emigration. The seasonal patterns are assumed to be between an even flow of net temporary migrants and the proportional flow as measured for permanent emigration which is expressed as an average of the two seasonal patterns. The mathematical expression that gives each monthly flow is as follows:

Equation 6.11:

$$NTE_m = \frac{\left[\left(\frac{NTE}{12}\right) + \left(\frac{EM_m}{EM} \times NTE\right)\right]}{2}$$

where

NTE_m = the number of net temporary emigrants for the month m;
NTE = the annual number of net temporary emigrants;
EM_m = the number of emigrants for the month m;
EM = the number of annual emigrants.

Net temporary emigration can only be estimated for the intercensal period preceding the most recent census. Postcensal estimates of net temporary emigration is assumed to be the same as those estimated in the previous intercensal period for each province and territory as no other source of information is available. They remain unchanged until the completion of the RRC in the next census (i.e., approximately two years after the census).

38. Special estimates are done for the territories as the RRC does not provide estimates.

Estimates of net temporary emigrants by age and sex

To obtain the estimates of net temporary emigrants by age and sex, the same distributions as observed for emigrant are used.[39]

Levels of estimates

The difference between preliminary and final estimates lies in the timeliness of the emigration estimate used to calculate the seasonal adjustment for the net temporary emigration. The same estimation method is used.

6.2.3 Returning emigrants

A returning emigrant is a person returning to Canada, after having been classified as an emigrant. In a manner similar to the procedure used to calculate the number of emigrants, data from the CCTB file and CRA's T1FF are used to estimate the number of returning emigrants. A citizen or permanent resident who has emigrated regains eligibility for the CCTB if he or she re-establishes residential, economic and social ties in Canada. A returning child emigrant is identified by the presence of both a departure date and return date on the CCTB file, as well as the parent's residency status. As with emigration, a person's return to Canada relies on their resident status for income tax purposes.

CCTB data for returning emigrants include children born outside Canada. These children are included in the returning emigrant population because the majority of foreign births to Canadian parents are not included in vital statistics (according to provincial registry offices). Their inclusion makes up for what would otherwise be an underestimation of population estimates for the reason of incompleteness of data on births. As is the case with emigrants, estimates of the number of returning emigrant children and the number of returning emigrant adults are calculated separately.

Returning emigrant children

The CCTB file provides numbers of dependent children (under the age of 18 years) of CCTB recipients who have returned to Canada after a period of emigration. These data are available on a monthly basis but provided by CRA to Statistics Canada annually for each province and territory.

As with emigrant children, adjustment factors are applied to compensate for the fact that the CCTB program is not universal. First, the CCTB data for returning child emigrants are adjusted with a factor reflecting the program's partial coverage, that is, people who do

39. Although the RRC provide age-sex data on temporary emigration, they are unreliable due to the large sampling variances at this level of detail.

apply for the CCTB or are not eligible. This factor is obtained by comparing the estimated number of children in the population with the number of children in CCTB files (Equation 6.2). The second adjustment factor is used to take into account the differential propensity to emigrate between children who are eligible for the CCTB and children who are not. In this case, it is assumed that the ratio is the same as the ratio for emigrants (Equation 6.3). Hence, the numbers for returning CCTB-eligible children are adjusted with the same differential propensity factors that are used for emigration. The formula provided earlier used to compute the estimate of child emigrants (Equation 6.1) is used to estimate returning child emigrants, where $({}^{CCTB}_{j}EM^{0-17})$ the number of child emigrants by province/territory, according to CCTB data, is replaced by $({}^{CCTB}_{j}RE^{0-17})$ the number of children returning to Canada.

Returning emigrant adults

The number of adults returning to Canada after emigrating can be indirectly estimated using ISD estimates based on income tax files or census data. Both sources include returning permanent emigrants, returning temporary emigrants including persons who may have maintained residential ties with Canada. However, neither source can be used directly to provide the number of returning adult permanent emigrants. It is possible to obtain the adult/child ratio in the returning emigrant population as follows:

Equation 6.12:

$$\tau^{RE} = \frac{{}^{C}RE^{18+}}{{}^{C}RE^{0-17}}$$

where

τ^{RE} = adult to child ratio of the returning emigrant population;
${}^{C}RE^{18+}$ = number of returning adult emigrants based on census counts;
${}^{C}RE^{0-17}$ = number of returning child emigrants based on census counts.

The number of returning adult emigrants is estimated by multiplying the estimate of returning child emigrants based on CCTB data by the adult/child ratio from Equation 6.12.

Equation 6.13:

$$RE^{18+} = {}^{CCTB}RE^{0-17} \times \tau^{RE}$$

where

RE^{18+} = estimated number of returning adult emigrants;
$^{CCTB}RE^{0-17}$ = estimate of returning child emigrants based on CCTB data;
τ^{RE} = adult/child ratio in the returning emigrant population based on census count.

The estimate of the number of returning children is produced monthly for each province and territory. The provincial/territorial distribution of returning children is assumed to hold for returning adults.

Estimates of returning emigrants by age and sex

The age and sex distribution of returning emigrants is based on Census of Canada data. Characteristics of returning emigrants are derived from census mobility data one year ago, after excluding non-permanent residents and immigrants. From 2006/2007 on, we use distribution by age and sex derived from 2006 Census. The census distribution by single years of age and sex is applied to children aged 0 to 17 years and to adults 18 years and over.

Levels of estimates

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. The same estimation method is used.

Chapter 7

Interprovincial migration

Movements of persons, who change their usual place of residence, crossing provincial/territorial boundaries, are classified as interprovincial migration. Intraprovincial migration (or subprovincial migration) also involves a change in usual place of residence, but these movements occur within the same province or territory. The term "internal migration" may refer to either interprovincial migration or intraprovincial migration, or both. This chapter focuses on the interprovincial migration component. For information on intraprovincial migration, see Chapter 8 on subprovincial estimates of population.

As is the case for emigration, there is no provision for recording interprovincial migration in Canada. Consequently, this component of demographic growth has to be estimated using administrative data. The methods used to prepare estimates of interprovincial migration will be discussed in this chapter.

7.1 Data sources and relevant concepts

Canada Revenue Agency (CRA) provides Statistics Canada with data from the Canada Child Tax Benefit (CCTB) program and personal income tax returns, which are used to estimate interprovincial migration. Preliminary[40] estimates by month are derived quarterly from CCTB data, while final estimates are derived annually from the T1 Family File (T1FF)[41].

Since 1976, personal income tax records have become the official data source for final interprovincial migration estimates. The population covered by tax data is more comprehensive than that of the CCTB; in general, interprovincial migration estimates based on tax data are considered to be of higher quality than those produced using data from the CCTB. However, tax data are not timely enough for use in preliminary estimates. Statistics Canada receives tax data only after annual processing, compared to CCTB data which are available on a monthly basis.

40. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.
41. The T1 Family File (T1FF) is derived from the Canada Revenue Agency (CRA) T1 file by Income Statistics Division of Statistics Canada.

7.2 Estimates of interprovincial migration, Canada, provinces and territories

7.2.1 Final estimates of interprovincial migration

Final estimates of interprovincial migration are produced using personal income tax data. By comparing the place of residence at the time of filing, it is possible to identify for those persons who file returns for two consecutive years, those who move, their place of origin and their destination. The development of migration data involves four main steps[42]:

1. geo-coding of tax records;
2. estimation of non-filing dependants of tax filers, by age groups and gender;
3. identification of migrant tax filers by age group and gender;
4. adjustment for the population not covered by the T1FF.

The four steps of the estimation methodology are described below.

Step 1—geo-coding

The migration data developed from the taxation records are estimates of migration flows between census divisions (CDs) or census metropolitan areas (CMAs). The geographic coding of CDs and CMAs on the tax records is done primarily on the basis of the postal code, which is a part of the mailing address. Tax records also contain a "locality code" assigned by Revenue Canada based on place name, which is a combination of Standard Geographic Classification (SGC) codes used by the census. About 99% of all tax records are filed with a postal code, or are assigned one, based on the filer's address by matching the address to a file similar to the Postal Code Directory from Canada Post. As the tax return is usually submitted several months after the end of the tax year, the postal codes correspond to those existing in the spring of the year following the tax year.

Step 2—estimation of dependants

Since the tax records source file has no direct information on the number and characteristics of non-filing dependants, this information must be imputed. Following the 1988 tax reforms, the estimation of tax filers' dependants is obtained from the T1FF. The family system creates families by linking all filing family members together, using spousal social insurance numbers, marital status, and matched addresses. This system imputes a non-filing spouse whenever a filer has declared himself/herself married but was not linked to a filing spouse. Children are imputed based on CCTB data, which carry the

42. For a detailed description of the methodology, see Statistics Canada 2010d.

social insurance number of the parent receiving the benefit. Newborns are added to tax filer records from Statistics Canada's vital statistics birth database, and in addition a historical file of imputed children is created from the previous year's tax file[43].

The family data from the tax file are then used to create a file of individual tax filers for the migration system. The migration file contains information for each tax filer and assigns dependants to each tax filer based on the number of filers in the family, and the number of dependants. If only one parent in the family has filed a tax return, then all dependants, including the spouse, if one exists, are added to the tax filer's record. If both parents filed tax returns, and their family record contained dependant children, then each tax filer would receive half the count of dependants.

If not already present, the age of the dependant spouse is estimated in the family system. The age of imputed children is provided (date of birth) on CCTB records, birth files from vital statistics, and from the Historical Family File. The family system does not impute gender for non-filing children. The gender for each dependant child is assigned a male/female ratio based on information from the most recent census.

Step 3—identifying migrant tax filers

After assigning the geography codes and non-filing dependants to each tax filer, records for two consecutive years are matched by social insurance number. Only records of tax filers who are present in both years are retained. Migrant tax filers are identified by comparing current and previous geography codes (CDs or CMAs). Thus, it is only possible to determine migrant status for those who file in two consecutive years. The identification of migrants is based on their address at the time of filing, and thus the migration period is not precisely one year, but it is assumed to be close. The assumed reference period is around April to April, of the year following each tax year. However, in the Demographic Estimates Program the given estimates cover the period from July to June. Tax filers' non-filing dependants are assumed to have the same migration behaviour as that of the filers to whom they are assigned.

Step 4—coverage adjustment factor

The final step in the estimation process is an adjustment for coverage, done at the CD/CMA level for gender and specific age groups. This adjustment is required to estimate those migrants who did not file two consecutive personal income tax returns, which involves the calculation of coverage adjustment factors (or inflation factors). Population estimates by CD/CMA are used as the denominator to the same population for whom two consecutive income tax returns could be linked to create coverage ratios.

43. The Historical Family File, updated annually by Income Statistics Division of Statistics Canada, consists of information of known family relationships obtained from tax returns.

For migration estimates prior and up to 2000/2001, provincial adjustment ratios were substituted in place of the CD/CMA ratio in the few cases where coverage was abnormally high or low (outlier values in the coverage ratios). Beginning with 2001/2002 migration data, high and low coverage cases were identified with a different methodology and a Canadian adjustment ratio was used in place of the CD/CMA ratio for these outlier coverage ratios. After a review and analysis of these methods, it was determined that cases where the coverage ratios were deemed outlier, represented small domains having little impact on the measure of migration, moreover, it was known that replacing the coverage ratios with a national ratio had the effect of reducing the true measure of migration flows. Consequently, beginning in 2006/2007, the identification and treatment of high and low coverage cases were discontinued and the coverage was simply adjusted to the Demography Division data at the CD level by age and sex group. The same adjustment is also done for CMAs.

The adjustment ratios are applied to the counts of interprovincial out-migrants derived in Step 3 to obtain an estimate of total migration. The basic assumption is that the population not covered by the taxation system has the same rate of migration as those for whom it can be measured.

7.2.2 Preliminary estimates of interprovincial migration

Preliminary estimates of migration between provinces and territories are produced using CCTB data, along with data from personal income tax records, processed by Income Statistics Division. Because the CCTB program is not universal and does not provide direct information on the number of adult migrants, this estimation requires three adjustment factors to derive the complete estimate of interprovincial migration:

1. the program's partial coverage (i.e., persons who do not apply to the CCTB program or who are not eligible);
2. the differential propensity to migrate between children who are receiving benefits from the CCTB and children who are not;
3. the differential propensity to migrate of adults as compared to that of children.

Each of these adjustment factors makes use of data from CCTB and T1FF. The estimation methods for child and adult interprovincial migration rates are described in the next section.

Interprovincial migrant children

Since income tax returns are annual and not available at the time when preliminary estimates are produced, the estimation of preliminary interprovincial migration is based on CCTB administrative files, which provide counts of child migrants (aged 0 to 17) receiving benefits under the program on a monthly basis. The data provide monthly information on the changes of usual residence involving provincial or territorial boundaries of CCTB-registered children receiving benefits and the total number of

children registered for the CCTB program by province and territory. As mentioned, the CCTB program is not universal and targets families with lower income or families with a larger number of dependent children; therefore the data are incomplete in terms of estimation of total child migrants and require adjustments.

The estimation of interprovincial migrant children involves two adjustment factors. The first factor reflects the CCTB program's coverage: the ratio of the number of children as estimated by Demography Division, to the number of children receiving benefits from the CCTB program, for each month and province and territory. The second factor, which corrects for any biases resulting from the differential propensity to migrate between children who are receiving benefits from the CCTB and children who are not, is obtained by comparing the interprovincial out-migration rates of children receiving the CCTB with the out-migration rates for all children (aged 0 to 17). The adjustment factor is calculated for each province and territory and is based on the two last available years of T1FF.[44]

The formula to estimate children migrating between provinces/territories is as follows:

Equation 7.1:

$$ {}_{j,k}\mathrm{IM}^{0-17} = {}_{j,k}\mathrm{CCTB\_IM}^{0-17} \times \frac{1}{{}^{\mathrm{CCTB}}_{j}\Re} \times {}_{j}\mathrm{G} $$

where

${}_{j,k}\mathrm{IM}^{0-17}$ = number of migrant children from province/territory of origin j to destination province/territory k;

${}_{j,k}\mathrm{CCTB\_IM}^{0-17}$ = interprovincial migration of children according to CCTB data by province/territory of origin j to destination province/territory k;

${}^{\mathrm{CCTB}}_{j}\Re$ = coverage rate of CCTB program for province/territory of origin j;

${}_{j}\mathrm{G}$ = adjustment factor for differential migration propensity among children for province/territory of origin j.

The equation above includes the adjustments for incomplete coverage of CCTB-children who receive benefits (Equation 7.2) and for the differences in the propensities of this group to migrate as compared to all children (Equation 7.3).

The coverage rates are calculated on a monthly basis as follows:

For each province and territory:

44. See Step 3, Section 7.2.1, for a description of the process to identify interprovincial migrants.

Equation 7.2:

$$ {}^{CCTB}_{j}\Re = \left[\frac{{}^{CCTB}_{j}P^{0-17}}{{}^{Dem}_{j}P^{0-17}} \right] $$

where

${}^{CCTB}_{j}\Re$ = coverage rate of CCTB program for each province/territory j;

${}^{CCTB}_{j}P^{0-17}$ = number of children aged 0 to 17 years registered for the CCTB program in province/territory j:

${}^{Dem}_{j}P^{0-17}$ = population estimated by Demography Division of children aged 0 to 17 years in province/territory j.

The differential migration propensity is based on a set of ${}_{j}G$ factors. Using tax data, the ${}_{j}G$ factors are obtained by dividing the interprovincial out-migration rates for all children by the out-migration rates of children receiving benefits from the CCTB program, as follows:

For each province and territory:

Equation 7.3:

$$ {}_{j}G = \frac{{}_{j}Rate\_Tax\_IM^{0-17}}{{}_{j}Rate\_CCTB\_IM^{0-17}} $$

$$ {}_{j}G = \frac{\dfrac{{}^{Tax}_{j}IM^{0-17}}{{}^{Tax}_{j}P^{0-17}}}{\dfrac{{}^{CCTB}_{j}IM^{0-17}}{{}^{CCTB}_{j}P^{0-17}}} $$

where

${}_{j}G$ = adjustment factor for the differential migration propensity by province/territory of origin j;

${}_{j}Rate\_Tax\_IM^{0-17}$ = interprovincial out-migration rate of all children from province/territory of origin j, according to income tax data;

${}_{j}Rate\_CCTB\_IM^{0-17}$ = interprovincial out-migration rate of CCTB-registered children from province/territory of origin j, as modeled from tax data (T1FF);

${}^{Tax}_{j}IM^{0-17}$ = number of total out-migrant children from province/territory of origin j, according to income tax data;

${}^{Tax}_{j}P^{0-17}$ = estimate of children in province/territory j based on income tax files;

${}^{CCTB}_{j}IM^{0-17}$ = number of out-migrant children eligible for CCTB benefits from province/territory j, according to tax data;

${}^{CCTB}_{j}P^{0-17}$ = CCTB-registered children in the population of the income tax files provided by T1FF, by province/territory j.

The ${}_{j}G$ factor will equal one if the migration rate of children receiving CCTB benefits and the migration rate of all children are identical. A value of ${}_{j}G$ that is greater (or less) than one indicates that the propensity to migrate of children receiving CCTB benefits is less (or greater) than that of all children.

Interprovincial migrant adults

The third factor in adjusting the estimates from the CCTB data is used to calculate the number of adult migrants. Estimates for adults migrating between provincial/territorial boundaries are obtained through the calculation of modeled migration ratios of adults to children referred to as F factors. The F factors, calculated from the most recent tax data, are equal to the ratio of the adult migration rate to the child migration rate, by province/territory of origin and destination province/territory, as follows:

For each province and territory of origin and destination:

Equation 7.4:

$$ {}_{j,k}F_{t-3,t} = \left[\frac{{}_{j,k}\text{AvgRate\_IM}^{18+}}{{}_{j,k}\text{AvgRate\_IM}^{0-17}} \right] $$

$$ {}_{j,k}F_{t-3,t} = \left[\frac{\dfrac{{}^{Tax}_{j,k}IM^{18+}_{t-3,t-2} + {}^{Tax}_{j,k}IM^{18+}_{t-2,t-1} + {}^{Tax}_{j,k}IM^{18+}_{t-1,t}}{{}^{Dem}_{j}P^{18+}_{t-2}}}{\dfrac{{}^{Tax}_{j,k}IM^{0-17}_{t-3,t-2} + {}^{Tax}_{j,k}IM^{0-17}_{t-2,t-1} + {}^{Tax}_{j,k}IM^{0-17}_{t-1,t}}{{}^{Dem}_{j}P^{0-17}_{t-2}}} \right] $$

where

$_{j,k}F_{t-3,t}$ = average adult estimation factor calculated over three migration periods by province/territory of origin j and province/territory of destination k;

$_{j,k}AvgRate\_IM^{18+}$ = average adult migration rate over three migration periods by province/territory of origin j and province/territory of destination k;

$_{j,k}AvgRate\_IM^{0-17}$ = average child migration rate over three migration periods by province/territory of origin j and province/territory of destination k;

$^{Tax}_{j,k}IM^{18+}$ = adult migration by province/territory of origin j and province/territory of destination k, according to income tax data for periods (t-3, t-2), (t-2, t-1), (t-1, t);

$^{Dem}_{j}P^{18+}_{t-2}$ = adult population estimated by Demography Division for province/territory of origin j at time (t-2);

$^{Tax}_{j}IM^{0-17}$ = child migration by province/territory of origin j and province/territory of destination k, according to income tax data for periods (t-3, t-2), (t-2, t-1), (t-1, t);

$^{Dem}_{j}P^{0-17}_{t-2}$ = child population estimated by Demography Division for province/territory of origin j at time (t-2).

The $_{j,k}F$ factors once applied to the child migration rates (based on adjusted CCTB data) produces estimates of adult inter-provincial/territorial migration rates. These are then multiplied by the provincial/territorial populations to obtain estimates of the number of adults migrating between provincial/territorial borders, as follows:

For each province and territory:

Equation 7.5:

$$_{j,k}IM^{18+} = \frac{_{j,k}IM^{0-17}}{^{Dem}_{j}P^{0-17}} \times {}^{Dem}_{j}P^{18+} \times {}_{j,k}F_{t-3,t}$$

where

$_{j,k}IM^{18+}$ = estimated number of adults migrating from province/territory of origin j to province/territory of destination k;

$\frac{_{j,k}IM^{0-17}}{^{Dem}_{j}P^{0-17}}$ = migration rate of children (estimates by province/territory of origin j and province/territory of destination k based on the adjusted CCTB data);

$^{Dem}_{j}P^{18+}$ = population estimate of adults at province/territory of origin j at the beginning of the period;

$_{jk}F_{t-3,t}$ = average adult estimation factor calculated over three migration periods by province/territory of origin j and province/territory of destination k.

Finally, after the estimates of the total number of interprovincial migrant children and adults have been separately calculated by province/territory of origin and destination, interprovincial migration estimates are obtained by adding the interprovincial migrants of children and adults, as follows:

For each province and territory:

Equation 7.6: $$_{jk}IM = {}_{jk}IM^{0-17} + {}_{jk}IM^{18+}$$

Interprovincial migrant population by age and sex

Final interprovincial migration estimates are derived from the T1FF file consisted solely by five broad age groups (under 18, 18 to 24, 25 to 44, 45 to 64, and 65 years and over) by sex, as well as origin and destination province/territory.

To estimate the number of interprovincial migrants by age and sex, the following steps are followed:

i. calculate migration by sex and five-year age group cohorts using the number of migrants by province of origin and destination obtained from the one-year ago mobility question from the most recent census;
ii. split the migration cohort into single-year of age using Sprague's multipliers;
iii. calculate a percentage distribution by single years of age for each broad age-sex group (under 18, 18 to 24, 25 to 44, 45 to 64, and 65 years and over), using the data obtained in step (ii);
iv. apply the distribution obtained in step (iii) (which remains constant for the five-year postcensal estimates) to the interprovincial migration estimates by broad age group and sex, in order to disaggregate them into a distribution by single-year of age;
v. Finally, to eliminate inconsistencies from one age to another, the number of migrants by single-year of age, is aggregated by five-year age groups, to which Sprague's multipliers are again applied.

To produce the preliminary interprovincial migration estimates by age and sex at provincial/territorial level, the most recent age and sex distribution of final estimates is applied.

Chapter 8

Subprovincial estimates of population

In addition to estimates at the national and provincial/territorial levels, population estimates are produced for subprovincial areas. Annual population estimates are released for census divisions (CDs), census metropolitan areas (CMAs), and economic regions (ERs). Historical series of population estimates at the subprovincial level is available from 1996 onwards. Custom requests for population estimates at other levels of geography (e.g., census subdivisions (CSDs), census agglomerations (CAs) and health regions) are also available.

This chapter discusses the estimation methods used for the production of postcensal and intercensal estimates of population for CDs, CMAs and ERs. The methods used to produce population estimates at census subdivision level are also presented at the end of the chapter.

8.1 Definition and relevant concepts

A **census division** (CD) is the general term for provincially legislated areas (such as county, *municipalité régionale de comté* and regional district) or their equivalents. Census divisions are intermediate geographic areas between the province and territory level and the municipality (i.e., census subdivision)[45]. Census divisions have been established in provincial law to facilitate regional planning, as well as the provision of services that can be more effectively delivered on a scale larger than a municipality. In Newfoundland and Labrador, Manitoba, Saskatchewan, Alberta, Yukon, Northwest Territories and Nunavut, provincial/territorial law does not provide for these administrative geographic areas. Therefore, Statistics Canada, in cooperation with these provinces and territories, has created equivalent areas called census divisions for the purpose of disseminating statistical data. In Yukon, the census division is equivalent to the entire territory.

A **census metropolitan area** (CMA) is formed by one or more adjacent municipalities centered on a large urban area (known as the urban core). A CMA must have a total population of at least 100,000 of which 50,000 or more must live in the urban core. To be included in the CMA, other adjacent municipalities must have a high degree of integration with the central urban area, as measured by commuting flows derived from census place of work data.

45. Census subdivision (CSD) is the general term for municipalities (as determined by provincial/territorial legislation) or areas treated as municipal equivalents for statistical purposes (e.g., Indian reserves, Indian settlements and unorganized territories).

Once an area becomes a CMA, it is retained as a CMA even if its total population declines below 100,000 or the population of its urban core falls below 50,000. The urban areas in the CMA that are not contiguous to the urban core are called the urban fringe. Rural areas in the CMA are called the rural fringe. All CMAs are subdivided into census tracts.

An **economic region** (ER) is a grouping of complete CDs (with one exception in Ontario) created as a standard geographic unit for analysis of regional economic activity. Within the province of Quebec, economic regions (*régions administratives*) are designated by law. In all other provinces and territories, economic regions are created by agreement between Statistics Canada and the province/territory concerned. Prince Edward Island and the three territories each consist of one economic region. In Ontario, there is one exception where the economic region boundary does not respect census division boundaries: the census division of Hamilton is split between the ER of Hamilton-Niagara Peninsula and the ER of Toronto.

8.2 Postcensal population estimates of subprovincial areas

8.2.1 Postcensal population estimates for CMAs and CDs

The component method is used to produce population estimates for CMAs and CDs, but figures are adjusted, if necessary, to assure consistency with provincial/territorial totals. Estimates of population are first produced for each province and territory, and then summed to obtain an estimate of the population of Canada. The estimates at the subprovincial level are adjusted proportionately to coincide with annual estimates of the total population for provinces and territories.

The component method formula used in estimating total CMA and CD populations is expressed as follows:

For every subprovincial area:

Equation 8.1:

$$P_{(t+i)} = P_t + B_{(t,t+i)} - D_{(t,t+i)} + I_{(t,t+i)} - (E_{(t,t+i)} + \Delta TE_{(t,t+i)}) + RE_{(t,t+i)} + \Delta NPR_{(t,t+i)} + \Delta NTER_{(t,t+i)} + \Delta NFRA_{(t,t+i)}$$

where, for each subprovincial area:

$(t,t+i)$ = interval between times t and t+i;

$P_{(t+i)}$ = estimate of the population at time t+i;

P_t = base population at time t (census adjusted for census net undercoverage or the most recent estimate);

$B_{(t,t+1)}$ = number of births;

$D_{(t,t+1)}$ = number of deaths;

$I_{(t,t+1)}$ = number of immigrants;

$E_{(t,t+1)}$ = number of emigrants;

$\Delta TE_{(t,t+1)}$ = net temporary emigration;

$RE_{(t,t+1)}$ = number of returning emigrants;

$\Delta NPR_{(t,t+1)}$ = net non-permanent residents;

$\Delta NTER_{(t,t+1)}$ = net interprovincial migration;

$\Delta NFRA_{(t,t+1)}$ = net subprovincial migration.

In the estimation of migration, an additional component, subprovincial migration or intraprovincial migration, (i.e., migration within provinces/territories and across subprovincial areas) is necessary.

8.2.2 Postcensal population estimates for economic regions (ERs)

A different method is used to produce population estimates for ERs: the census division (CD) aggregation method. First, the ERs are defined in terms of CDs using Standard Geographical Classification (SGC) specifications. When the geographic delineation of the CDs matches that of the ER, no adjustment is required; the population estimates for the CDs that make up the ER are simply added together.

However, when the geographic delineation of the CDs does not match that of the ER, i.e., when a CD is in more than one ER, allocation of the CD's demographic components is prorated on the basis of its proportion of each ER's population. The proportions are referred to as *conversion factors*. They are calculated using the most recent census counts.

Subsequently, all demographic components, with the exception of intraprovincial migration, initially measured at the CD level can be allocated to each ER. In other words, the population and demographic components of ERs can be estimated by aggregating the CD data based on the ERs geographic delineation.

The CD aggregation method is not an appropriate method to estimate the components of intraprovincial migration for ERs as it overestimates the number of in-migrants and out-migrants. In-migrants to a given CD from another CD in the same ER should not be counted as the migration occurred within the ERs boundaries. Thus, they are false in-migrants. The same is true for out-migrants from one CD to another CD in the same ER. Again, they are false out-migrants. However, combining the in-migration and out-migration figures as derived by the CD aggregation method produces a consistent result since the false in-migrants and false out-migrants cancel out. Thus, only the net intraprovincial migration of ERs can be estimated accurately by the CD aggregation

method, and due to the reasons stated above, the numbers of intraprovincial in-migrants and out-migrants are not available for ERs.

Special treatment for preliminary postcensal estimates for Quebec and British Columbia

A different method is used to calculate preliminary postcensal population estimates for CDs and CMAs in Quebec and British Columbia. For Quebec, the total population estimates produced by the *Institut de la statistique du Québec* (ISQ) are used. Those estimates are based on data from the insured persons *Fichier d'inscription des personnes assurées* (FIPA) of the *Régie de l'assurance-maladie du Québec* (RAMQ). These estimates are adjusted to Demography Division's estimate of Quebec's total population. The same approach is followed for the ISQ's ER-level estimates, which are derived from its CD-level estimates.

For British Columbia, preliminary postcensal estimates are computed using CD and CMA growth rates provided by British Columbia's Statistical Agency (BC STATS), for total population only. To produce new population estimates for each CD and CMA, the rates are applied to Demography Division's estimates of total population for the previous year. Prorating is used to ensure that the province's total population released by Demography Division equals the sum of the populations of the subprovincial areas. The British Columbia population estimates used to calculate the growth rates are produced using the Difference-Correlation method, a regression model based on residential electrical (Hydro) connections and *Ministry of Health Client Registry* data as symptomatic indicators.

8.2.3 Subprovincial postcensal population estimates by age and sex

Postcensal population estimates by age and sex for CDs and CMAs are produced by applying the component method to each age-sex cohort in the base population, whereby the population is aged from year to year and the components are tabulated according to age and sex cohorts. Chapter 9 describes the application of the cohort component approach in details.

The component-method formula for estimating postcensal population estimates by age and sex of CMAs and CDs are as follows:

For age 0:

Equation 8.2:

$$P^{0}_{(t+1)} = B_{(t,t+1)} - D^{-1}_{(t,t+1)} + I^{-1}_{(t,t+1)} - \left[E^{-1}_{(t,t+1)} + \Delta TE^{-1}_{(t,t+1)}\right] + RE^{-1}_{(t,t+1)} + \Delta NPR^{0}_{(t,t+1)} + \Delta NTER^{-1}_{(t,t+1)} + \Delta NFRA^{-1}_{(t,t+1)}$$

Ages 1 to 89:

Equation 8.3:

$$P^{a+1}_{(t+1)} = P^{a}_{(t)} - D^{a}_{(t,t+1)} + I^{a}_{(t,t+1)} - \left[E^{a}_{(t,t+1)} + \Delta TE^{a}_{(t,t+1)}\right] + RE^{a}_{(t,t+1)} + \Delta NPR^{a}_{(t,t+1)} + \Delta NTER^{a}_{(t,t+1)} + \Delta NFRA^{a}_{(t,t+1)}$$

For age 90 and over:

Equation 8.4:

$$P^{90+}_{(t+1)} = P^{89+}_{(t)} - D^{89+}_{(t,t+1)} + I^{89+}_{(t,t+1)} - \left[E^{89+}_{(t,t+1)} + \Delta TE^{89+}_{(t,t+1)}\right] + RE^{89+}_{(t,t+1)} + \Delta NPR^{89+}_{(t,t+1)} + \Delta NTER^{89+}_{(t,t+1)} + \Delta NFRA^{89+}_{(t,t+1)}$$

where, for each subprovincial region:

$(t,t+1)$ = interval between times t and t+1;
$P_{(t+1)}$ = population estimates at time t+1;
$P_{(t)}$ = base population at time t (census adjusted for census net undercoverage or the most recent estimate);
B = number of births;
D = number of deaths;
I = number of immigrants;
E = number of emigrants;
ΔTE = net temporary emigrants;
RE = number of returning emigrants;
ΔNPR = net non-permanent residents;
$\Delta NTER$ = net interprovincial migration;
$\Delta NFRA$ = net subprovincial migration.

In order to ensure the concordance between the subprovincial and provincial and territorial population estimates by age and sex, we use the prorating technique. This method ensures the coherence between subprovincial and provincial/territorial population estimates for each age and sex combination.

Special treatment for preliminary postcensal estimates by age and sex for Quebec and British Columbia

A different method is used to calculate preliminary postcensal population estimates by age and sex for CDs, CMAs and ERs in Quebec. Same as the preliminary total population estimation, the population estimates by age and sex produced by the Institut de la statistique du Québec (ISQ) are used. These estimates are based on data from the

insured persons *Fichier d'inscription des personnes assures* (FIPA) of the *Régie de l'assurance-maladie du Québec* (RAMQ). They are benchmarked to correspond to Demography Division's total estimates by age and sex for Quebec.

In the case of British Columbia, preliminary postcensal estimates at the CMA and CD level are first obtained at the total population level as described in the previous section. Afterward, the preliminary postcensal estimates are split by age and sex using the age and sex distribution obtained by the component method of Demography Division. The same approach is followed for the ER-level estimates, which are derived from the CD-level estimates.

In order to ensure the concordance between the subprovincial and provincial estimates by age and sex, a prorating technique is used. This method ensures the coherence between subprovincial and provincial population estimates for each age and sex combination.

8.2.4 Levels of estimates

For Quebec and British Columbia, methods described in the previous sections are used only for preliminary postcensal population estimates. For updated and final postcensal estimates, the component method is used.

For the other provinces and territories, the difference between preliminary[46] and final postcensal population estimates lies in the timeliness of the components. When all the components are preliminary, the population estimate is described as preliminary postcensal (PP). When they are all final, the estimate is referred to as final postcensal (PD). Any other combination of levels is referred as updated postcensal (PR).

8.2.5 Base population and components of population growth

Base population

A full description of the methodology for deriving the postcensal base population is described in Chapter 2. In the Demographic Estimates Program, the base populations for subprovincial areas are derived from the quinquennial censuses between 1996 and 2006. Population counts at both the provincial/territorial and subprovincial levels are subjected to the same adjustment procedures as outlined in Chapter 2, unless otherwise noted. To estimate census net undercoverage (CNU) at the subprovincial level, provincial and territorial CNU rates by age and sex were applied to all geographic regions (census metropolitan areas (CMAs) and census divisions (CDs) in the province).

Prior to generating the population estimates for each component, the total province/territorial population counts and the counts for CDs and CMAs are adjusted, if necessary, to assure consistency between the two sets of figures. To adjust these data,

46. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.

two adjustment procedures are used: prorating and two-way raking[47]. Prorating is used to ensure that the province's total population equals the sum of the populations of the subprovincial areas. The technique adjusts the total population counts to the components of subprovincial areas. It distributes the difference between the province's total population and the sum of the population of the subprovincial areas proportionally across those areas.

Two-way raking is an iterative procedure whereby differences are simultaneously adjusted in assuming a consistency between:

1. the sum of subprovincial areas and provincial/territorial population totals, and
2. the sum across subprovincial areas for specific age/sex categories and the provincial/territorial distribution by age and sex.

Since CD and CMA boundaries do not remain stable over time, component data are adjusted to respect the boundaries defined in the 2006 Census. This ensures a stable base population for all estimates.

Births and deaths

In addition to the national and provincial/territorial data on births and deaths, the numbers of births and deaths at the census divisions (CDs) as for the deaths for the census metropolitan areas (CMAs) are derived directly from the vital statistics database of Statistics Canada's Health Statistics Division. The births for CMAs will also be calculated the same way starting in 2007/2008. A different method was used to produce estimates of births for 2006/2007 called the census division conversion method. According to this method, each CMA is first defined in terms of CDs (complete or in part), using the Standard Geographical Classification (SGC) specifications.

In cases when the geographic delineation of a group of CDs matches that of a CMA, no adjustment is required; the births of each CD that make up the CMA are simply added together. However, when the geographic delineation of the CD does not match that of the CMA, i.e., when a CD is not totally included in a CMA, a proportion of the CD's births is allocated to the CMA. This proportion is equal to the fraction of the CD's population living in the CMA territory.

These proportions are referred to as *conversion factors*. They are calculated using the most recent census counts. With these factors and the census division conversion method, it is possible to allocate to each CMA, the births initially measured at the CD level.

47. The two-way raking method is also referred to as the "Deming method", the "method of iterative proportions", and calibration (see Shryock, Siegel et al. 1976).

Estimates of births and deaths are categorized as final when they are directly taken from Health Statistics Division's vital statistics. To ensure their consistency, the estimates are subsequently controlled to the provincial totals using two-way raking.

When no data are available for births and deaths, subprovincial estimates are produced by distributing the preliminary provincial/territorial estimates on the basis of the most recent subprovincial distributions derived from Health Statistics Division's vital statistics. In such cases, the estimates of births and deaths are categorized as preliminary. To ensure their consistency, the estimates are subsequently controlled to the provincial/territorial totals using two-way raking.

Quebec and British Columbia provide their most recent estimates of births and deaths. These estimates are used for the preliminary estimates. However, the final estimates of births and deaths for these provinces are derived directly from the vital statistics database of Statistics Canada's Health Statistics Division.

Immigration

The distribution of the number of immigrants by CD and CMA cannot be obtained from the Citizenship and Immigration Canada's (CIC's) files as with the provincial/territorial level estimates, because the geographic coding is specific to CIC and is not directly convertible to the coding used by Statistics Canada. Subprovincial immigration estimates are produced by disaggregating the preliminary and final provincial/territorial estimates on the basis of the most recent subprovincial distribution derived from T1 Family File (T1FF)[48]. The data are available only by broad age groups (0 to 17, 18 to 24, 25 to 44, 45 to 64, and 65 years and over) and must be disaggregated by sex and single years of age based on the provincial distribution. To ensure their consistency, they are subsequently controlled to the provincial/territorial totals using two-way raking.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the estimates of the number of immigrants are derived from provincial/territorial estimates, the level of the subprovincial estimates will be the same. Immigration estimates are preliminary the first year and updated the following year. They become final two years after the reference year.

Net non-permanent residents (NPRs)

At the subprovincial level, there are no reliable administrative data available to estimate NPRs. To compensate for the lack of data, the provincial/territorial estimates by age and sex are disaggregated by subprovincial area on the basis of the subprovincial distribution in the most recent census. To ensure their consistency, they are subsequently controlled to the provincial/territorial totals using two-way raking.

48. The T1 Family File (T1FF) is derived from the Canada Revenue Agency (CRA) T1 file by Income Statistics Division of Statistics Canada.

Since the estimates of the net number of NPRs are derived from provincial/territorial estimates, the level of the subprovincial estimates will be the same. Estimates of NPRs are preliminary the first year and updated the following year. They become final two to three years after the reference year.

Emigration

As in the case of immigrants, the number of emigrants at the subprovincial level is derived from the T1FF. The estimates are available only by broad age groups (0 to 17, 18 to 24, 25 to 44, 45 to 64, and 65 years and over) and must be disaggregated by sex and single years of age based on the provincial/territorial distribution. To ensure their consistency, they are subsequently controlled to the provincial/territorial totals using two-way raking.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the estimates of the number of emigrants are derived from provincial/territorial estimates, the level of the subprovincial estimates will be the same.

Net temporary emigration

At the subprovincial level, provincial/territorial net temporary emigration is disaggregated on the basis of the age and sex distribution of subprovincial emigrants. To ensure their consistency, the estimates are subsequently controlled to the provincial/territorial totals using two-way raking.

The difference between preliminary and final estimates lies in the timeliness of the estimate of the net temporary emigration.

Returning emigrants

In order to estimate the numbers of returning emigrants at the subprovincial level, provincial/territorial numbers of returning emigrants are disaggregated on the basis of the age and sex distribution of subprovincial immigrants. To ensure their consistency, they are subsequently controlled to the provincial/territorial totals using two-way raking.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the estimates of the number of returning emigrants are derived from provincial/territorial estimates, the level of the subprovincial estimates will be the same.

Interprovincial migration

Interprovincial migration by broad age group and sex for subprovincial areas is derived from the T1FF. The estimates by broad age group and sex are disaggregated into single years of age using distributions from the 2006 Census one-year mobility question. To ensure their consistency, they are subsequently controlled to the provincial/territorial totals using two-way raking.

The difference between preliminary and final estimates lies in the timeliness of the sources used to estimate this component. Since the estimates of the number of migrants are derived from provincial/territorial estimates, the level of the subprovincial estimates will be the same.

Intraprovincial migration

As in the case of interprovincial migration, the components of intraprovincial migration by broad age group and sex are derived from the T1FF for both preliminary and final estimates. The estimates by broad age group and sex are disaggregated into single years of age using distributions from the 2006 Census one-year mobility question.

The difference between preliminary and final estimates lies in the timeliness of the T1FF data used to estimate this component.

Given the fact that there are no reliable data for preliminary intraprovincial migration estimates, the data of the most recent year, for which final estimates are available, are used. It is assumed that intraprovincial migratory behaviours of the current year are similar to those of the last year for which final estimates are available.

8.3 Intercensal population estimates of subprovincial areas

Intercensal estimates for the 1996 to 2006 period at the census division (CD), census metropolitan area (CMA), and economic region (ER) levels are produced much in the same manner as intercensal estimates at the provincial/territorial level (for information on the methods, refer to Chapter 1). There are three main steps in the production of intercensal estimates:

- the correspondence of the boundaries between both censuses;
- calculation of the error of closure;
- linear distribution of the error of closure.

However, the production of these estimates for CDs, CMAs and ERs additionally require that changes in census geography over time are accounted for. To ensure the correspondence of the boundaries for the 1996 to 2006 period, the base population and components of population growth from 1996 to 2006 had to be adjusted for the 2006

Census boundaries. For areas whose boundaries changed between the two censuses (2001 and 2006 Standard Geographical Classification (SGC)), *conversion factors* based on the 2006 Census subdivisions were used. In general, the corrections to CDs, CMAs, and ERs are minor.

As for the provinces/territories, the error of closure is defined as the difference between the postcensal population estimates on Census Day and the population enumerated in that census adjusted for census net undercoverage (CNU). The error of closure is spread uniformly over the intercensal period. Intercensal estimates by age and sex are adjusted in the same way, i.e., by distributing the error of closure uniformly across the age and sex cohorts. Like the postcensal estimates, the subprovincial intercensal estimates by age and sex are adjusted to ensure consistency with the provincial estimates using two-way raking.

8.4 Population estimates for census subdivision

A **census subdivision** (CSD) is the general term for municipalities (as determined by provincial/territorial legislation) or areas treated as municipal equivalents for statistical purposes (e.g., Indian reserves, Indian settlements and unorganized territories). They are classified into 55 types according to official designations adopted by provincial/territorial or federal authorities. Two exceptions are Subdivision of unorganized (SNO) in Newfoundland and Labrador, and Subdivision of county municipality (SC) in Nova Scotia, which are geographic areas created as equivalents for municipalities by Statistics Canada, in cooperation with those provinces, for the purpose of disseminating statistical data.

Census subdivision populations are not estimated using the demographic components of growth (births, deaths and migration), as it is the case with estimates for census divisions (CDs), census metropolitan areas (CMAs), provinces and territories. For CSDs, there are no demographic components that are sufficiently accurate. Accordingly, the alternate method that is used applies the census division (CD) growth rate to the base population of a CSD. For this method to be used, two types of data are necessary: the base population of the CSDs and the annual population estimates for the CDs.

8.4.1 Base population of census subdivision

The adjusted census counts for each CSD by age and sex constitute the base population. The base populations are derived from the quinquennial censuses between 1996 and 2006[49] and these base populations are adjusted as follows:

- adjustment to take population reviews into account. Because there are no population reviews by age and sex, calibration by age and sex is carried out to ensure consistency with total counts for population reviews;

49. The population universe of the 2006 Census is discussed in Chapter 2.

- adjustment of the population for census net undercoverage (CNU). Given that CNU rates at the subprovincial level are not sufficiently accurate, provincial and territorial rates by age and sex are used;
- addition of independent population estimates for incompletely enumerated Indian reserves in 1996, 2001 and 2006; and adjustment for early enumeration in 1996[50].

8.4.2 Postcensal population estimates for census subdivision

Census subdivision populations are estimated in three steps: first, the census division (CD) populations are estimated; second, the growth rate for the CDs are calculated by age and sex; and third, the growth rates for the CDs are applied to the CSDs that comprise them.

Step 1—Estimating CD populations

The first step is to estimate the population of the CDs using the component method as described previously in this chapter. This method is used to produce population estimates at the CD level by age and sex.

To ensure consistency between subprovincial estimates and provincial and territorial estimates by age and sex, a prorating technique is used for calibration.

Step 2—Calculating CD growth rates by age and sex

The second step consists of calculating the growth rate for the CDs by age and sex. The formula is as follows:

Equation 8.5:

$$GR_{\_j}CD^{a}_{(t,t+1)} = \frac{Pop_{\_j}CD^{a}_{(t+1)} - Pop_{\_j}CD^{a}_{(t)}}{Pop_{\_j}CD^{a}_{(t)}}$$

where,

$(t,t+1)$ = interval between time t and t+1;

$GR_{\_j}CD^{a}_{(t,t+1)}$ = growth rate of CD j at age a for period t and t+1;

$Pop_{\_j}CD^{a}_{(t)}$ = population of CD j at age a and time t.

50. There was no adjustment for early enumeration in 2001 and 2006.

Step 3—postcensal CSD population estimates

The final step consists of applying the growth rate by age and sex for a CD to the base population of the CSDs that comprise it. The formula is as follows:

Equation 8.6:

$$Pop_{\_iCDjCSD^{a}_{(t+1)}} = Pop_{iCDjCSD^{a}_{(t)}} + \left[Pop_{\_iCDjCSD^{a}_{(t)}} \times GR_{\_jCD^{a}_{(t,t+1)}}\right]$$

where

$Pop_{\_iCDjCSD^{a}_{(t)}}$ = population of CSD i of CD j at age a and time t;

$GR_{\_jCD^{a}_{(t,t+1)}}$ = growth rate of CD j at age a for period t and t+1.

Finally, in order to ensure consistency between CSD estimates and CD estimates by age and sex, prorating technique is used for calibration.

Chapter 9

Population estimates by age, sex and marital status

Analysis of population by age and sex is a fundamental aspect of most demographic studies. The age and sex structure of the population varies with time and place, while at the same time demographic behaviour is often a function of age and sex. For example, mortality rates are much higher in the older age groups. High migration rates are associated with young adults, as they move for personal and economic reasons. Population estimates by age and sex are widely used by other divisions of Statistics Canada. For example, these estimates are used in the calculation of employment and unemployment rates and crime rates, which tend to vary according to age and sex distributions. The addition of marital status builds the foundation for studying other demographic phenomena such as marriage and divorce rates and changes in family structure. Government and private sector planning and policies are largely driven by the age, sex and marital status profiles of certain populations.

Population estimates by age and sex are available at national, provincial/territorial, census division, census metropolitan area and economic region levels. The more detailed breakdown of these estimates by marital status is available only at the national and provincial/territorial levels. This chapter presents the methods used to produce population estimates disaggregated by age, sex and marital status.

9.1 Postcensal population estimates by age and sex, Canada, provinces and territories

9.1.1 Data sources and relevant concepts

Postcensal estimates of population by age and sex are produced using the cohort component approach. This is similar to the component method as used in the production of total population estimates, although additional data are required in its application. The data required for the cohort component method are related to demographic events (deaths, immigration, net non-permanent residents, emigration, returning emigration, net temporary emigration and interprovincial migration) that can be directly linked to persons belonging to the same birth cohort (i.e., persons having been born during the same period or year). Different components require unique treatment, according to the nature of the data used to generate the estimates, and their respective chapters elaborate upon the manner in which the estimate for each component is distributed by age and sex.

The data sources used in the production of the population estimates by age and sex are as follows[51]:

- Births and deaths using vital statistics;
- Immigration and non-permanent residents using data from Citizenship and Immigration Canada (CIC);
- Emigration are distributed by age and sex using data from the Office of Immigration Statistics, U.S. Department of Homeland Security, and Income Statistics Division (ISD);
- Net temporary emigration are distributed by age and sex using emigration distributions;
- Returning emigrants are distributed by age and sex using the most recent census data on mobility data one year ago, after excluding non-permanent residents and immigrants;
- Interprovincial migration by age and sex is derived from T1 family file by Income Statistics Division and counts from the last available census (one-year mobility variable).

9.1.2 Levels of estimates

The difference between preliminary[52] and final postcensal estimates lies in the timeliness of the components. When all the components are preliminary, the estimate is described as preliminary postcensal (PP). When they are all final, the estimate is referred to as final postcensal (PD). Any other combination of levels is referred to as updated postcensal (PR).

9.1.3 Methods of estimation

Postcensal estimates of population by age and sex are produced using the cohort component approach. This approach requires a slight modification of the component approach described in Chapter 1, but the overall principles are the same.

Annual estimates

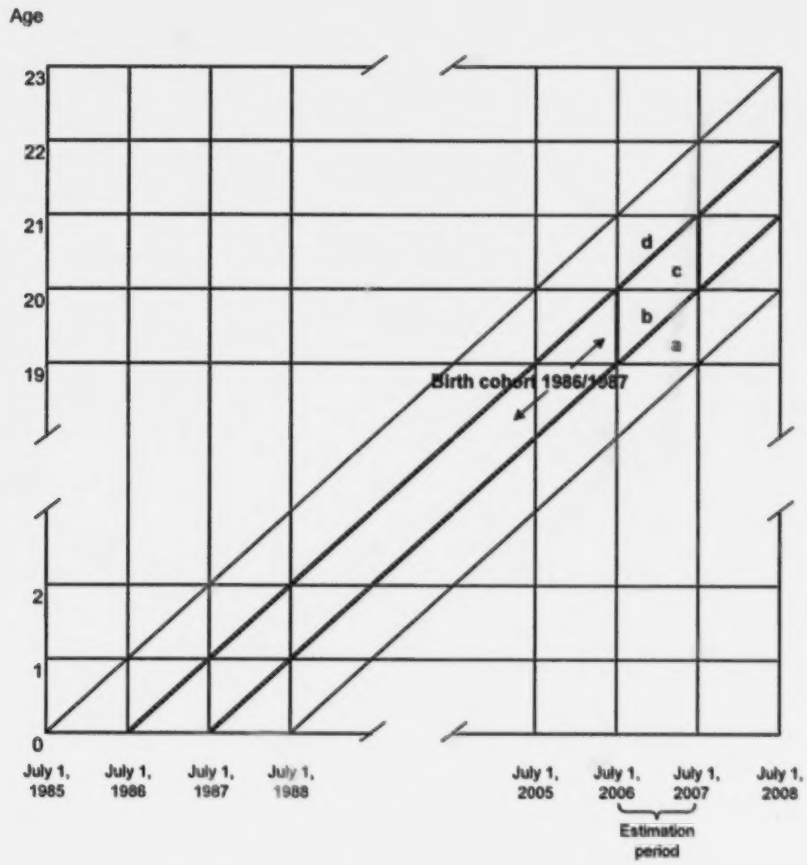
Estimates of population by age and sex are published annually with July 1st as the reference date. To calculate these estimates, birth cohorts (those persons born during the same year) for both males and females separately, are used. The cohort component approach factors in the aging of the cohorts over time. For example, persons aged 19 one year will be 20 years old the following year. The data required for the cohort component method include demographic events such as deaths, immigration, emigration, net

51. For the methodology used to produce estimates by age and sex for each component of demographic growth, refer to their respective chapters in this report.
52. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.

temporary emigration, returning emigration, non-permanent residents and interprovincial migration that can be directly linked to persons belonging to the same birth and sex cohorts.

Demographers use a tool called a Lexis diagram (Figure 9.1) to aid in the linking of events to specific cohorts. Time is located on the horizontal axis (abscissa), while the vertical axis (ordinate) represents age. Specific cohorts are identified by the diagonals (lifelines) that cross the diagram. Using the cohort component approach, demographic events are organized to follow these lifelines.

Figure 9.1
Transition from a distribution of demographic events by age and period to a distribution by age and birth cohort



Take, for example, those aged 19 as of July 1, 2006, who belong to the cohort born between July 1, 1986 and June 30, 1987 (inclusive). The demographic events

experienced by this cohort during the estimation period are represented by triangles "b" and "c".

The equations for estimating annual population by single years of age and sex, by the cohort component method (at the national and provincial/territorial levels) are as follows:

For each sex, by province/territory:

At age 0[53]:
Equation 9.1:

$$P^{0}_{(t+1)} = B_{(t,t+1)} - D^{-1}_{(t,t+1)} + I^{-1}_{(t,t+1)} - (E^{-1}_{(t,t+1)} + \Delta TE^{-1}_{(t,t+1)}) + RE^{-1}_{(t,t+1)} + \Delta NPR^{-1}_{(t,t+1)} + \Delta N^{-1}_{(t,t+1)}$$

From 1 to 99 years:
Equation 9.2:

$$P^{(a+1)}_{(t+1)} = P^{a}_{t} - D^{a}_{(t,t+1)} + I^{a}_{(t,t+1)} - (E^{a}_{(t,t+1)} + \Delta TE^{a}_{(t,t+1)}) + RE^{a}_{(t,t+1)} + \Delta NPR^{a}_{(t,t+1)} + \Delta N^{a}_{(t,t+1)}$$

For 100 years and over:
Equation 9.3:

$$P^{100+}_{(t+1)} = P^{99+}_{t} - D^{99+}_{(t,t+1)} + I^{99+}_{(t,t+1)} - (E^{99+}_{(t,t+1)} + \Delta TE^{99+}_{(t,t+1)}) + RE^{99+}_{(t,t+1)} + \Delta NPR^{99+}_{(t,t+1)} + \Delta N^{99+}_{(t,t+1)}$$

where

$(t, t+1)$	=	interval between times t and t+1;
a	=	age;
$P_{(t+1)}$	=	estimate of the population at time t+1;
P_t	=	base population at time t (census adjusted for CNU[54], or most recent estimate);
B	=	number of births;
D	=	number of deaths;
I	=	number of immigrants;
E	=	number of emigrants;
ΔTE	=	net temporary emigration;
RE	=	number of returning emigrants;
ΔNPR	=	net non-permanent residents;
ΔN	=	net interprovincial migration.

53. Age -1 refers to the cohort who had a demographic event at age 0 during the interval (t, t+1) but who was not born at the beginning of the interval.
54. Unless otherwise noted, the adjustment for the census net undercoverage (CNU) also includes the incompletely enumerated Indian reserves.

Annual population estimates by single years of age and sex for persons aged 0 to 99, and 100 years and over are available from 2001 and onwards at the national, provincial and territorial level. Previously, the upper limit of the age category was age 90 and over.

9.2 Intercensal population estimates by age and sex, Canada, provinces and territories

Intercensal population estimates for reference dates between two censuses are produced following each census. They reconcile previous postcensal estimates with the new census counts. Like the total population by province/territory, intercensal population by age and sex are adjusted by distributing the error of closure uniformly across the age-sex cohorts.

9.3 Subprovincial postcensal and intercensal estimates by age and sex

Postcensal population estimates by age and sex for census divisions (CDs) and census metropolitan areas (CMAs) are produced by applying the component method to each age-sex cohort in the base population, whereby the population is aged from year to year and the components are tabulated according to age and sex cohorts. A different method called the census division (CD) aggregation method is used to produce population estimates by age and sex for economic regions (ERs). Descriptions of the methods used to estimate the populations by age and sex for CMAs, CDs and ERs are provided in Chapter 8. At the subprovincial level, annual population estimates by age and sex are available for ages 0 to 89 and age 90 and over.

Special methods for preliminary postcensal estimates by age and sex are applied for CDs, CMAs and ERs in Quebec and British Columbia. These methods and the approach used to derive intercensal estimates by age and sex at subprovincial levels are described in Chapter 8.

9.4 Population estimates by age, sex and marital status, Canada, provinces and territories

There are two series of population estimates by marital status, the main difference between them being the treatment of persons living in "common law" unions. One of them is the series of estimates by legal marital status, i.e., a person's conjugal status under the law (for example, single, married, widowed or divorced). On the basis of this definition, people living "common law" are categorized by their legal marital status. If a person has never "married" and is living "common law", he or she is regarded as "single" under this definition.

The other is the series of estimates by marital status, i.e., a person's "de facto" conjugal status. For example, a person who reports being legally "widowed" and is living with

another person as a couple but is not married to that person will be counted as "common law" in the marital status series and "widowed" in the legal marital status series.

Separate estimates for legal and "de facto" marital statuses at the national, provincial and territorial levels are available from 1991 onwards. However, estimates for the "de facto" marital status exist since 1971. Estimates of marital statuses are not produced for subprovincial levels.

9.4.1 Data sources and relevant concepts

Marital status refers to the conjugal status of a person. In demographic estimates, a distinction is made between "legal" marital status and marital status. The distinction between the two definitions lies in the concept of who is considered married. In the discussion of legal marital status, a person's marital status is determined by law. Common-law partners are not legally married to each other, thus are considered single, divorced or widowed according to their legal marital status. Separated couples are considered married under both concepts.

The following definitions represent those used by Statistics Canada for legal marital status and marital status, respectively.

Legal marital status refers to the marital status of the person under the law. Estimates are presented in the following categories: single, legally married, widowed or divorced.

Single: Persons who have never been married or persons whose marriage has been annulled and have not remarried. All persons aged less than 15 are considered as never married (single);

Legally married: Persons whose spouse is living, unless a divorce has been obtained. Persons separated are also included in this category;

Widowed: Persons who have lost their spouse through death and who have not remarried;

Divorced: Persons who have obtained a divorce and who have not remarried.

Marital status indicates the conjugal arrangement of a person. Estimates are presented in the following categories: single, married (including persons living common-law and persons who are separated), widowed or divorced. Common-law status refers to whether the person aged 15 or over is living with a person of the opposite sex or of the same sex as a couple but is not legally married to that person. It includes situations where the members of such a couple are living apart temporarily because of illness, work or school.

Single: Persons who have never been married, or persons whose marriage has been annulled and who have not remarried, and do not live common-law. All persons aged less than 15 are considered as never married (single);

Married (including persons living common-law and persons who are separated): Persons whose spouse is living, unless a divorce has been obtained;

Widowed: Persons who have lost their spouse through death and who have not remarried, and who do not live common-law;

Divorced: Persons who have obtained a divorce and who have not remarried, and who do not live common-law.

To produce population estimates by age, sex and marital status, data about events resulting in a change of marital status (for example, marriage, divorce, death of a spouse) are also required. These data are available from the vital statistics database maintained by Health Statistics Division of Statistics Canada. The database was created in collaboration with provincial and territorial ministries as well as the federal department of Justice Canada. This collaboration allows Statistics Canada to obtain data on marriage certificates registered by the provinces/territories for legal purposes as well as data related to divorces from the Central Divorce Registry[55] maintained by Justice Canada.

9.5 Postcensal population estimates by legal marital status

9.5.1 Methods of estimation

Population estimates by legal marital status are produced by the cohort component method. This method takes into account events that result in a change of legal marital status. All persons under the age of 15 years are assumed to be single. For persons aged 15 years and over, estimates by single year of age are prepared for four subpopulations: single, legally married, divorced and widowed, as previously defined for legal marital status. These estimates are based on the census counts, adjusted to July 1st, as well as for census net undercoverage.

The component method formulae used to estimate the four legal marital status subpopulations are as follows:

For each sex and age (15 years and older), by province/territory[56]:

Single:
Equation 9.4:

$$P_{(t+1)}^{nm,(a+1)} = P_t^{nm,a} - D_{(t,t+1)}^{nm,a} + I_{(t,t+1)}^{nm,a} - (E_{(t,t+1)}^{nm,a} + \Delta TE_{(t,t+1)}^{nm,a}) + RE_{(t,t+1)}^{nm,a} + \Delta NPR_{(t,t+1)}^{nm,a} + \Delta N_{(t,t+1)}^{nm,a} - Mar_{(t,t+1)}^{nm,a}$$

55. The Central Divorce Registry is a database containing all the information related to divorce decrees granted in Canadian courts.
56. All parts of Equation 9.4 to Equation 9.7 refer to legal marital status.

Legally married:
Equation 9.5:

$$P_{(t+1)}^{mar,(a+1)} = P_t^{mar,a} - D_{(t,t+1)}^{mar,a} + I_{(t,t+1)}^{mar,a} - (E_{(t,t+1)}^{mar,a} + \Delta TE_{(t,t+1)}^{mar,a}) + RE_{(t,t+1)}^{mar,a} + \Delta NPR_{(t,t+1)}^{mar,a} + \Delta N_{(t,t+1)}^{mar,a} + Mar_{(t,t+1)}^{a} - Veu_{(t,t+1)}^{a} - Div_{(t,t+1)}^{a}$$

Divorced:
Equation 9.6:

$$P_{(t+1)}^{v,(a+1)} = P_t^{v,a} - D_{(t,t+1)}^{v,a} + I_{(t,t+1)}^{v,a} - (E_{(t,t+1)}^{v,a} + \Delta TE_{(t,t+1)}^{v,a}) + RE_{(t,t+1)}^{v,a} + \Delta NPR_{(t,t+1)}^{v,a} + \Delta N_{(t,t+1)}^{v,a} - Mar_{(t,t+1)}^{v,a} + Div_{(t,t+1)}^{a}$$

Widowed:
Equation 9.7:

$$P_{(t+1)}^{w,(a+1)} = P_t^{w,a} - D_{(t,t+1)}^{w,a} + I_{(t,t+1)}^{w,a} - (E_{(t,t+1)}^{w,a} + \Delta TE_{(t,t+1)}^{w,a}) + RE_{(t,t+1)}^{w,a} + \Delta NPR_{(t,t+1)}^{w,a} + \Delta N_{(t,t+1)}^{w,a} - Mar_{(t,t+1)}^{w,a} + Veu_{(t,t+1)}^{a}$$

where for each province and territory:

a	=	age;
(t,t+1)	=	interval between times t and t+1;
nm	=	never married;
mar	=	legally married;
v	=	divorced;
w	=	widowed;
P_{t+1}^{a+1}	=	estimate of the population at age a+1 at time t+1;
P_t^a	=	base population at time t (census adjusted for CNU, or most recent estimates) at age a;
B	=	number of births;
D	=	number of deaths;
I	=	number of immigrants;
E	=	number of emigrants;
ΔTE	=	net temporary emigration;
RE	=	number of returning emigrants;
ΔNPR	=	net non-permanent residents;
ΔN	=	net interprovincial migration;
Mar	=	number of marriages;
Div	=	number of divorces;
Veu	=	new widowhoods.

The components of population growth are disaggregated by legal marital status as follows:

1. **deaths** by age and sex are disaggregated by legal marital status using the latest death statistics published by Health Statistics Division;
2. **interprovincial migration** is disaggregated by legal marital status using the counts from the census used in calculating the base population (one-year mobility question);
3. **immigration and non-permanent residents** are disaggregated by legal marital status using the estimates derived from the files of Citizenship and Immigration Canada's Field Operational Support System (FOSS);
4. **emigration and net temporary emigration** is disaggregated by legal marital status using the latest available data from the Office of Immigration Statistics, U.S. Department of Homeland Security;
5. **number of returning emigrants** is disaggregated by legal marital status using the population counts from the census used in the base population (one-year mobility question);
6. **marriages** are disaggregated by age, sex and legal marital status using the latest marriage statistics published by Health Statistics Division. The totals are estimated beforehand by Demography Division;
7. **divorces** are disaggregated by age, sex and legal marital status using the latest Justice Canada divorce statistics published by Health Statistics Division. The totals are estimated beforehand by Demography Division;
8. the difficulty with deaths of married persons is that the age of the surviving spouse is not recorded in the Vital Statistics Registry. Consequently, the age of the new widow or widower must be determined indirectly. **New widowhoods** are derived from estimates of deaths of married persons and husband-wife families based on the population counts from the census used in calculating the base population.

New widowhoods

Using the age distribution of deaths to married males and females, compiled by age[57] group, the surviving spouse is assigned to a five-year age group. Assignment is based on the distribution of legally married husband-wife families by age group of husband and age group of wife from the last census (see Table 9.1).

57. Age at the beginning of the estimation period.

Table 9.1
Percent distribution of husband-wife families by age group of wife and age group of husband according to the legal marital status concept: Ontario, 2001 Census

Age group of husband (years)	Age group of wife (years)								
	15 to 19 (1)	20 to 24 (2)	25 to 29 (3)	30 to 34 (4)	35 to 39 (5)	40 to 44 (6)	45 to 49 (7)	50 to 54 (8)	55 to 59 (9)
15 to 19	19.0	13.6	12.2	5.7	14.2	8.9	12.8	2.0	0.8
20 to 24	3.4	65.4	23.8	3.6	0.9	0.6	0.9	0.7	0.1
25 to 29	0.3	17.7	64.1	15.1	2.0	0.4	0.2	0.1	0.1
30 to 34	0.1	3.3	28.6	53.2	12.1	2.1	0.4	0.1	0
35 to 39	0	0.7	6.1	31.6	49.5	10.0	1.6	0.4	0.1
40 to 44	0	0.2	1.5	7.8	35.5	45.4	8.0	1.3	0.2
45 to 49	0	0	0.3	1.9	9.5	37.1	43.0	7.1	0.9
50 to 54	0	0	0.1	0.6	2.6	10.9	37.3	41.8	5.6
55 to 59	0	0	0.1	0.3	0.9	3.1	11.5	41.7	36.3
60 to 64	0	0	0	0.2	0.4	1.1	3.3	13.7	40.5
65 to 69	0	0	0	0.1	0.2	0.5	1.2	4.3	12.8
70 to 74	0	0	0	0.1	0.1	0.2	0.5	1.5	4.3
75 to 79	0	0	0	0	0.1	0.1	0.2	0.6	1.4
80 to 84	0	0.1	0	0	0	0.1	0.1	0.4	0.5
85 to 89	0	0	0	0.1	0.1	0.2	0.1	0.2	0.3
90 and over	0.2	0	0.7	0.5	1.0	0.5	0	0	0.4

	60 to 64 (10)	65 to 69 (11)	70 to 74 (12)	75 to 79 (13)	80 to 84 (14)	85 to 89 (15)	90 and over (16)	Total (17)
15 to 19	1.4	3.9	1.4	4.1	0	0	0	100.0
20 to 24	0.3	0	0.1	0.1	0.1	0	0	100.0
25 to 29	0	0	0	0	0	0	0	100.0
30 to 34	0.1	0	0	0	0	0	0	100.0
35 to 39	0	0	0	0	0	0	0	100.0
40 to 44	0.1	0	0	0	0	0	0	100.0
45 to 49	0.2	0	0	0	0	0	0	100.0
50 to 54	0.9	0.2	0	0	0	0	0	100.0
55 to 59	5.2	0.7	0.2	0	0	0	0	100.0
60 to 64	34.4	5.3	0.8	0.2	0	0	0	100.0
65 to 69	38.9	34.9	5.9	1.1	0.1	0	0	100.0
70 to 74	13.4	36.5	35.7	6.7	0.9	0.1	0	100.0
75 to 79	3.9	13.3	39.9	35.0	5.0	0.4	0.1	100.0
80 to 84	1.8	4.9	17.8	42.3	28.1	3.6	0.3	100.0
85 to 89	1.0	2.2	8.5	25.3	38.7	21.1	2.2	100.0
90 and over	0.5	0.6	5.2	12.2	24.3	36.3	17.6	100.0

Note: The percentages in this table are rounded to sum exactly to 100.0.
Source: 2001 Census data.

The number of new widow(er)s by age group (a, $a + 4$), for $a = 15, 20,\ldots,85$, is estimated as follows for each province/territory:

Equation 9.8:

$$\mathrm{Veu}^{(a,a+4)} = \sum_{(b=15)}^{90+} \mathrm{D}^{mar,(b,b+4)} \times \Phi^{(a,a+4)}$$

with

Equation 9.9:

$$\Phi^{(a,a+4)} = \frac{\mathrm{F}^{r(a,a+4)}}{\sum_{(a=15)}^{90+} \mathrm{F}^{r(a,a+4)}}$$

where

$Veu^{(a,a+4)}$	=	estimated number of new widow(er)s in age group (a, $a + 4$);
$D^{mar,(b,b+4)}$	=	number of deaths of married persons (male or female) in age group (b, $b + 4$);
$\Phi^{(a,a+4)}$	=	proportion of husband-wife families by five-year age group of husbands (wives) cross-classified by five-year age group of wives (husbands) (a, $a + 4$);
$F'^{(a,a+4)}$	=	number of husband-wife families by five-year age group of husbands (wives) cross-classified by five-year age group of wives (husbands) (a, $a + 4$).

Table 9.2 provides an illustration of the above formulae. A sample calculation of the incidence of widowhood among females in the 45 to 49 year age group for the province of Ontario in 2001/2002 is presented. The percent distribution of husbands with a 45 to 49 wife is taken from column (7) in Table 9.1. This is applied to the 2001/2002 distribution of male deaths (column (1) in Table 9.2), and the results summed to estimate the number of new widows aged 45 to 49 years, as presented in column (3). To determine the number of widows in another age group, the appropriate distribution from Table 9.1 is substituted for column (2) of Table 9.2, and the same procedure is followed.

Table 9.2
Calculation of the incidence of widowhood according to the legal marital status concept: An example for females aged 45 to 49 years, Ontario, 2001/2002

Age group of husband (years)	Deaths of married males	Percent distribution of husbands with wives aged 45 to 49 years	Estimates of new widows aged 45 to 49 years
	(1)	(2)	(3) = [(1) X (2)] / 100
15 to 19	0	12.8	0.000
20 to 24	10	0.9	0.086
25 to 29	54	0.2	0.091
30 to 34	143	0.4	0.601
35 to 39	226	1.6	3.699
40 to 44	430	8.0	34.189
45 to 49	695	42.9	298.397
50 to 54	1,146	37.3	427.597
55 to 59	1,474	11.5	169.526
60 to 64	2,099	3.3	68.605
65 to 69	2,907	1.2	35.143
70 to 74	4,039	0.5	18.732
75 to 79	4,613	0.2	10.373
80 to 84	3,864	0.1	3.982
85 to 89	2,402	0.1	1.766
90 and over	960	0.0	0.258
Total			**1,073.045**
			Rounded to: 1,073

Source: Demography Division, Statistics Canada.

These results by five-year age group are then disaggregated into single years of age using Sprague's multipliers[58]. This distribution of surviving spouses represents the spouse's age at the beginning of the reference period.

9.5.2 Levels of estimates

The difference between preliminary and final postcensal estimates of the population by legal marital status lies in the timeliness of the components. When all the components are preliminary, the estimate is described as preliminary postcensal (PP). When they are all final, the estimate is referred to as final postcensal (PD). Any other combination of levels is referred to as updated postcensal (PR).

58. Sprague's multipliers are interpolation coefficients used to subdivide data. For a detailed description of the use of Sprague's multipliers, see Shryock, Siegel et al. 1976.

9.6 Postcensal population estimates by marital status

9.6.1 Methods of estimation

Since there are no reliable sources of data on annual formations and dissolutions of common-law unions, an alternative method is used to produce postcensal estimates by marital status. The number of persons living common-law is estimated using estimates by legal marital status and proportions derived from the census adjusted for census net undercoverage. The proportion of never-married persons living common-law is obtained by dividing the number of never-married persons living common-law by the total number of never-married persons. The same procedure is used to calculate the proportion of divorced and widowed persons living common-law. The marital status proportions are generated for each age, sex, province and territory.

The derived proportions are then multiplied by the number of never-married, divorced and widowed persons respectively to produce the number of persons living common-law. The number of persons living common-law by marital status is then subtracted from the estimates by legal marital status and added to the married category to produce the population by marital status.

As for the estimates by age and sex, the base populations by marital status are adjusted to July 1st and for census net undercoverage.

Postcensal population estimates by marital status for a given age and sex can be expressed as follows:

For each province/territory, by sex:

Single:
Equation 9.10:

$$P^{nm(ms),a}_{(t+1)} = P^{nm(legal),a}_{(t+1)} - P^{nm(cl),a}_{(t+1)}$$

Married:
Equation 9.11:

$$P^{mar(ms),a}_{(t+1)} = P^{mar(legal),a}_{(t+1)} + P^{nm(cl),a}_{(t+1)} + P^{w(cl),a}_{(t+1)} + P^{v(cl),a}_{(t+1)}$$

Divorced:
Equation 9.12:

$$P^{v(ms),a}_{(t+1)} = P^{v(legal),a}_{(t+1)} - P^{v(cl),a}_{(t+1)}$$

Widowed:

Equation 9.13:

$$P^{w(ms),a}_{(t+1)} = P^{w(legal),a}_{(t+1)} - P^{w(cl),a}_{(t+1)}$$

where

Equation 9.14:

$$P^{nm(cl)}_{(t+1)} = P^{nm(legal)}_{(t+1)} \times \frac{P^{nm(cl)}_{c}}{P^{nm(legal)}_{c}}$$

Equation 9.15:

$$P^{v(cl)}_{(t+1)} = P^{v(legal)}_{(t+1)} \times \frac{P^{v(cl)}_{c}}{P^{v(legal)}_{c}}$$

Equation 9.16:

$$P^{w(cl)}_{(t+1)} = P^{w(legal)}_{(t+1)} \times \frac{P^{w(cl)}_{c}}{P^{w(legal)}_{c}}$$

where

(t+1)	=	at time (t+1);
a	=	age;
$P^{nm(ms),a}_{(t+1)}$	=	population of single persons according to the marital status concept;
$P^{nm(legal),a}_{(t+1)}$	=	population of single persons according to the legal marital status concept;
$P^{nm(cl),a}_{(t+1)}$	=	population of persons living in common-law unions, never previously married;
$P^{mar(ms),a}_{(t+1)}$	=	population of married persons according to the marital status concept;
$P^{mar(legal),a}_{(t+1)}$	=	population of married persons according to the legal marital status concept;
$P^{w(cl),a}_{(t+1)}$	=	population of persons living in common-law unions who were widowed by their previous spouse;
$P^{v(cl),a}_{(t+1)}$	=	population of persons living in common-law unions who divorced their previous spouse;
$P^{v(ms),a}_{(t+1)}$	=	population of divorced persons according to the marital status concept;

$P_{(t+1)}^{v(legal),a}$ = population of divorced persons according to the legal marital status concept;

$P_{(t+1)}^{w(ms),a}$ = population of widowed persons according to the marital status concept;

$P_{(t+1)}^{w(legal),a}$ = population of widowed persons according to the legal marital status concept;

$P_c^{nm(cl)}$ = censal estimate of population of single persons living in common-law unions;

$P_c^{nm(legal)}$ = censal estimate of population of single persons according to legal marital status concept;

$P_c^{v(cl)}$ = censal estimate of population of divorced persons living in common-law unions;

$P_c^{v(legal)}$ = censal estimate of population of divorced persons according to the legal marital status concept;

$P_c^{w(cl)}$ = censal estimate of population of widowed persons living in common-law unions;

$P_c^{w(legal)}$ = censal estimate of population of widowed persons according to the legal marital status concept.

The two series of population estimates by marital status sum to the corresponding age and sex estimates of population by province and territory.

9.6.2 Levels of estimate

The difference between preliminary and final postcensal estimates stems from the estimates by legal marital status. The same estimation method is used in both sets of estimates. As a result, the estimates by marital status have the same revision level as the estimates by legal marital status from which they are derived.

9.7 Intercensal population estimates by marital status and legal marital status

The production of intercensal estimates by age and sex is done by distributing the error of closure across age and sex cohorts. For a description on the calculation and distribution of the error of closure, see Chapter 1.

Adjusted census distributions by age, sex and marital status/legal marital status from the two most recent censuses are used to derive intercensal estimates of population by marital status/legal marital status. The census distributions are linearly interpolated to obtain the required series of distributions. The interpolated distributions are then applied to the

intercensal population estimates by age and sex to obtain estimates by age, sex and marital status/legal marital status.

9.8 Changes to the methodology

The methodology presented in this chapter was used until the release of the July 1, 2007 estimates. A new method was used for the following releases.

Given the fact that marriage and divorce data are no longer available, Demography Division had to modify the estimation method to produce population estimates by marital status and legal marital status.

Since marital status patterns do not really change over a short period, Demography Division has opted for the method of census ratios. At each census, a series of ratios by age, sex, marital status and legal marital status will be produced. The ratios will remain stable for the entire postcensal period and will be applied to the annual population estimates by age and sex.

Intercensal estimates will continue to be produced according to the method presented in this chapter.

Chapter 10

Estimates of census families, economic entities and households

In addition to population estimates, Statistics Canada produces annual estimates of census families for Canada, provinces and territories. These estimates are available from 1986 onwards. Estimates of the number of census families can be either postcensal or intercensal. Except for the territories, estimates of census families are available by family type, sex and age group of parents, family size and children's age group.

Due to a reorganization of the Demographic Estimates Program, the Demography Division has used a case-specific method to produce postcensal estimates for the 2007 to 2009 period. This is different from the components method[59] that was used for the previous years and the headship rates method (using income tax files) that is implemented in 2011. The following discussion on the estimation method will focus on the headship rates method and readers can refer to the previous version of this report on details of the components method. The interim case-specific method will also be discussed in this chapter, as the changes in methods affect the consistency of the historical series of estimates.

The Demographic Estimates Program also produces estimates of economic entities and private households at the national and provincial/territorial levels by the headship rates method. Estimates for these two analytical units are used internally[60] and are available on request. In this chapter, reference to the development of private households and economic entities is also presented as harmonized estimation of these three demographic indicators can be attained with the new headship rates method.

10.1 Definitions and relevant concepts

For the purpose of generating these estimates, the definitions of the three concepts are the same as those in the 2006 Census[61].

A **census family** refers to a married couple (with or without children of either or both spouses), a couple living common-law (with or without children of either or both partners) or a lone parent of any marital status, with at least one child living in the same

59. For details on the components method, refer to Statistics Canada 2007b.

60. These estimates are used in the calculation of demographic, social and economic indicators. They are used for planning, program evaluation and base population for various surveys and studies. In addition, these data are also used in calculating of weights for use in Statistics Canada surveys.

61. For more details, see the 2006 Census Dictionary (http://www12.statcan.gc.ca/census-recensement/2006/ref/dict/index-eng.cfm).

dwelling. A couple may be of opposite or same sex[62]. "Children" in a census family include grandchildren living with their grandparent(s) but with no parents present. For the estimates, married couples and couples living common-law are grouped into a single category.

An **economic entity** is derived from the concept of economic family. It includes economic families and persons not living in an economic family (who accordingly constitute economic entities of size 1). According to the census, an economic family is defined as a group of two or more persons who live in the same dwelling and are related to each other by blood, marriage, common-law or adoption. A couple may be of opposite or same sex. For 2006, foster children are included.

A **household** refers to a person or a group of persons (other than foreign residents) who occupy the same dwelling and do not have a usual place of residence elsewhere in Canada. It may consist of a family group (census family) with or without other persons, of two or more families sharing a dwelling, of a group of unrelated persons, or of one person living alone. Household members who are temporarily absent on Census Day (e.g., residing temporarily elsewhere) are considered as part of their usual household. For census purposes, every person is a member of one and only one household. Unless otherwise specified, all data in household reports are for private households only.

10.2 Data sources

Four sources of statistical data, all described previously in this document, are needed to produce estimates of these three entities: tax data from the T1 Family File (T1FF), census data, data on undercoverage from coverage studies, and population estimates. Table 10.1 shows the sources, references and role of the data used for estimating these analytical units.

62. Before the implementation of the headship rates method in 2011, same-sex couples were excluded in the estimation of census families, since it was impossible to establish trends with regard to the increase of this group by the components method.

Table 10.1
Sources, references and role of data used to produce estimates of census families, economic entities and households

Source	Reference	Role
T1 Family File	Income Statistics Division, Statistics Canada	Creation of census correction coefficients
Census	Statistics Canada	Creation of census correction coefficients Distribution of size of analytical units of 6 or more persons
Census coverage studies	Social Survey Methods Division, Statistics Canada	Adjustment of census net undercoverage
Population estimates	Demography Division, Statistics Canada	Adjustment of reference date

10.3 Postcensal estimates of census families, economic entities and households

Annual estimates of census families, economic entities and private households are based on T1FF tax data with different adjustments. Even though the T1FF data do not create economic entities and private households, it can be used to estimate these analytical units. The method can be divided into three steps: correct the biases inherent in the T1FF, adjust the reference date and adjust the census coverage.

Step 1—correction of biases inherent in the T1FF

The coverage provided by T1FF tax data, while excellent, is not complete. Also, some operational definitions in the T1FF do not directly correspond to census definitions. Therefore, it is necessary to make corrections to the T1FF by comparing the number of heads of census families in the T1FF with the number in the census for each of the three analytical units. The way of determining heads varies depending on the concept:

- Census families:
  - Couples: each spouse shares the role of family head and is therefore equivalent to half a head;
  - Single-parent families: the parent is the head;
  - Non-family: each individual is a head.
- Economic entities:
  - Couples: each spouse is equivalent to half a head, as in the case of census families;
  - Single-parent families: the parent is the head;

  - Non-family: each individual is a head.
- Private Households:
  - The primary household maintainer is the head (a household may have more than one maintainer).

It is therefore possible to create a census correction coefficient for each combination of characteristics, namely province/territory of residence, age group, sex, type of census family, age group of spouse, sex of spouse, size of census family and age group of children. The way to calculate this coefficient depends on the situation, as Table 10.2 shows.

Table 10.2
Possible situations for calculating census correction coefficients

Situation	$H_{CF}^{T1FF}(p,a,s,fam,a',s',sf,a'')$	$H_i^{Cen}(X)$	$E_i(X)$	Level
1	>0	>0	$\frac{H_i^{Cen}(X)}{H_{CF}^{T1FF}(p,a,s,fam,a',s',sf,a'')}$	8
2	=0	=0	0	0
3	>0	=0	0	0
4	=0	>0	See below	

where

E = Census correction coefficient is a ratio of census heads of census families, economic entities and households by characteristics (X) to T1FF heads of census families by (p,a,s,fam,a',s',sf,a'')

sf = size of census family
H^{T1FF} = T1FF heads
H^{Cen} = census heads
X = p,a,s,fam,a′,s′,sf,a″,efam,se,sh
p = province/territory of residence
a = age group
s = sex
fam = type of census family
a′ = age group of spouse
s′ = sex of spouse
sh = size of household
a″ = age group of children
efam = type of economic family
se = size of economic family
i = CF, EE, or HSD
CF = census families
EE = economic entities
HSD = households

In Table 10.2, 'Level' refers to census family characteristics. For example, in the first situation, the correction coefficient is available for each combination of the eight characteristics. In the second and third situations, a correction coefficient of 0 indicates no estimates of heads in the census.

The correction coefficient in the fourth situation is used to estimate economic entities and private households with T1FF. These census correction coefficients are calculated each census year and are assumed to remain constant for the entire postcensal period. They represent both an adjustment of coverage and an adjustment of conceptual bias which is illustrated in the following decomposition:

Equation 10.1:

$$E_i(X) = \text{bias} \times \text{relationship} = \frac{H_i^{Cen}}{H_i^{T1FF}} \times \frac{H_i^{T1FF}}{H_{CF}^{T1FF}} = \frac{H_i^{Cen}}{H_{CF}^{T1FF}}$$

Each year, it will be possible to correct the biases inherent in the T1FF simply by multiplying the number of heads in the T1FF by the census correction coefficients, taking the level of each coefficient into account.

Equation 10.2:

$$CH_i(X) = E_i(X) \times H_{CF}^{T1FF}(X)$$

where

$CH_i(X)$ is the corrected number of heads in the T1FF.

Step 2—adjustment of the reference date

The T1FF is based on tax data that are mostly collected no later than the end of April and often refer to the situation on December 31, whereas the demographic estimates generally have different reference dates (such as July 1). The census correction coefficient in Step 1 takes into account the growth between December 31 and Census Day. Here, the correction applied to the T1FF data set adjusts the reference date from Census Day of the current year to July 1. This time lag is factored in using headship rates, constructed with population estimates. Since census net undercoverage will be taken into account in the next step, the population estimates used for the headship rates exclude the adjustment for census net undercoverage. For an estimate at July 1, the adjustment of the reference date is as follows:

Equation 10.3:

$$CH_i^{Jul1}(X) = \frac{CH_i^{CD}(X)}{\text{Unadjust\_pop}_{CD}^{Dem}(p,a,s)} \times \text{Unadjust\_pop}_{Jul1}^{Dem}(p,a,s)$$

where

$CH_i^{Jull}(X)$ = corrected number of heads in the T1FF at July 1;

$CH_i^{CD}(X)$ = corrected number of heads in the T1FF at Census Day;

$Unadjust\_pop_{CD}^{Dem}(p,a,s)$ = $Pop_{CD}^{Dem}(p,a,s) - Pop\_CNU(p,a,s)$;

$Unadjust\_pop_{Jull}^{Dem}(p,a,s$ = $Pop_{Jull}^{Dem}(p,a,s) - Pop\_CNU(p,a,s)$;

Pop_CNU = population obtained on the basis of census net undercoverage;

CD = Census Day of current year.

Step 3—adjustment of census coverage

Since the census omits some individuals, it is logical that this also results in the omission of census families, economic entities and private households. A correction is therefore needed to take these omitted analytical units into account.

For census families, the adjustment method involves calculating the probability of an individual being omitted in the census in one of the following strata, according to his or her age group and de facto marital status:

- adult in a couple category with no children
- adult in a couple category with children
- single parent
- child
- adult in a non-family household
- adult in a collective dwelling.

These probabilities are then applied to the figures for census net undercoverage by province/territory, age, sex and marital status to obtain the number of heads of census families. Finally, census weights are adjusted to include these new census families by characteristics, province/territory, sex, age group, de facto marital status, and household size. For economic entities and private households, a ratio between census family heads and economic entities and household heads is calculated for each combination of characteristics. This ratio is then applied to the estimates of census net undercoverage of census families to obtain the estimate of census net undercoverage of economic entities and households. These estimates of census net undercoverage are held constant for the entire postcensal period.

The new methodology proposed for 2006 census net undercoverage of families was modified due to a particular issue in the 2006 Reverse Record Check (RRC). The information used to ascertain the probabilities that a missed person was in a particular family stratum, given their demographic characteristics, was not available in a timely manner to produce the estimates. For that reason, the probabilities were based on data taken from the 2001 RRC findings.

Final equation

By combining all three steps above, it becomes possible to construct the final equation for producing estimates. For an estimate at July 1, the equation is:

Equation 10.4:

$$EST_i(X) = \frac{H_{CF}^{T1FF}(X) \times E_i(X)}{Unadjust\_pop_{CD}^{Dem}(p,a,s)} \times Unadjust\_pop_{Jul1}^{Dem}(p,a,s) + EST\_CNU_i(X)$$

where

$EST_i(X)$ = estimated number of census families, economic entities or households;

$EST\_CNU_i(X)$ = estimate of census net undercoverage of census families, economic entities or households.

10.4 Characteristics of census families, economic entities and households

As seen in the previous section, estimates are produced for several characteristics. These are estimated at the same time as the number of analytical units. These characteristics are:

- province/territory of residence
- broad age group of heads
- sex of head(s)
- type of census family (couple, single-parent family, non-census family)
- size of census family (2, 3, 4, 5, 6 and more members)
- broad age group of children
- type of economic entity (couple, single-parent family, non-economic family)
- size of economic entity (2, 3, 4, 5, 6 and more members)
- size of household (2, 3, 4, 5, 6 and more members)

10.5 Postcensal estimates of number of persons living in analytical units

In addition to producing estimates of the number of census families, economic entities and households, it is also possible to produce an estimate of the number of persons associated with each of these analytical units. This estimate is produced in two parts: first, for units of which the size is 1 to 5 persons and second, for units of size 6 and more persons.

Size of 1 to 5 persons

For these units, the estimate is direct, since all that is needed is to multiply the size ($Size_i(X)$) by the number of units ($EST_i(X|Size \in [1,5])$) to obtain the number of persons in each analytical unit with size of 1 to 5 persons, denoted as ($IND_i(X|Size \in [1,5])$):

Equation 10.5:

$$IND_i(X|Size \in [1,5]) = EST_i(X|Size \in [1,5]) \times Size_i(X)$$

Size of 6 and more persons

For these units, since their actual size is suppressed in the construction of the model, we use provincial/territorial distributions from the last census to estimate the average size for analytical units of 6 and more persons, here denoted as ($AverageSize_i(p)$), and then apply it to the estimates of each analytical unit ($EST_i(X|Size = 6+)$). The number of persons in each analytical unit with size of 6 and more persons can be obtained by using the following equation:

Equation 10.6:

$$IND_i(X|Size = 6+) = EST_i(X|Size = 6+) \times AverageSize_i(p)$$

Estimated number of persons

We can obtain the total number of persons living in each analytical unit by summing the two results obtained.

Equation 10.7:

$$IND_i(X) = IND_i(X|Size \in [1,5]) + IND_i(X|Size = 6+)$$

10.6 Level of estimates

Postcensal estimates may be preliminary[63] or final; the difference being in the level of timeliness of source files.

- Preliminary postcensal estimates:

63. Unless otherwise noted, the term *preliminary* includes both preliminary and updated estimates.

  - Preliminary population estimates;
  - T1FF for the fiscal year two years prior to the reference year (for example, the T1FF for the 2012 fiscal year for estimates at July 1, 2014);
- Final postcensal estimates:
  - Final population estimates;
  - T1FF for the fiscal year preceding the reference year (for example, the T1FF for the 2013 fiscal year for estimates at July 1, 2014).

10.7 Intercensal estimates of census families

Once the postcensal cycle is completed, intercensal estimates can be calculated. The method remains the same as in components method; only the parameters of the model are changed. The census correction coefficients, the average sizes for the units consisting of 6 or more persons and the units added to adjust for census net undercoverage are interpolated linearly, while the population estimates used are intercensal

10.8 Previous methods of estimating census families

In the past, census families were estimated with the components method (for the number of census families) or by extrapolation (for characteristics), while households and economic entities were estimated by the headship rates method. Since those methods have limitations, they were replaced by the method (headship rates method using income tax files) described in this chapter. It is therefore important to note that estimates released before 2010 or with a reference date of 2005 and prior were produced using the old methods. Through the use of the headship rates method based on income tax files, same-sex couples are included in the estimation of census families.

Due to a reorganization of the Demography Estimates Program, the Demography Division has used a case-specific method to produce reliable postcensal census family estimates for the 2007 to 2009 period. The main source of information used in this method is the census (for the years 2001 and 2006) adjusted for census net undercoverage. The postcensal estimates for the total number of census families by province and territory, and for the family characteristics in the case of the provinces only, are produced using a methodology that comprises two steps: 1) in order to produce the postcensal estimates, the increase (in percent) observed between 2005 and 2006 was applied to the estimate of the number of census families for the year 2006 at 2008. This calculation was applied to the provinces and territories; 2) total number of families estimated in the previous step is distributed according to family type, sex and age of parents, family size and children's age by extrapolating changes in these distributions as observed between the two previous censuses.

Appendix 1

Glossary

Age

Age as of July 1.

Census coverage

Census net undercoverage (CNU): Difference between undercoverage and overcoverage.

Overcoverage: Number of persons or families who should not have been counted in the census or who were counted more than once.

Undercoverage: Number of persons or families who were intended to have been enumerated in a census but were not enumerated.

Census division (CD)

Census division (CD) is the general term for provincially legislated areas (such as county, municipalité régionale de comté and regional district) or their equivalents. Census divisions are intermediate geographic areas between the province level and the municipality (census subdivision).

In Newfoundland and Labrador, Manitoba, Saskatchewan, Alberta, Yukon, Northwest Territories and Nunavut, provincial/territorial law does not provide for these administrative geographic areas. Therefore, Statistics Canada in cooperation with these provinces and territories has created equivalent areas called census divisions for disseminating statistical data. In Yukon, the census division is equivalent to the entire territory.

Census family

Refers to a married couple (with or without children of either or both spouses), a couple living common-law (with or without children of either or both partners) or a lone parent (of any marital status) with at least one child living in the same dwelling. A couple may be of opposite or same sex. The "children" in a census family include the grandchild who live in the household of at least one of their grandparents, in the absence of the parents.

Census metropolitan area (CMA)

A census metropolitan area (CMA) is formed by one or more adjacent municipalities centred on a large urban area (known as the urban core). A CMA must have a total population of at least 100,000 of which 50,000 or more must live in the urban core. To be included in the CMA, other adjacent municipalities must have a high degree of integration with the central urban area, as measured by commuting flows derived from census place of work data.

Once an area becomes a CMA, it is retained as a CMA even if the population declines below 100,000 or its urban core population declines below 50,000. The urban areas in the CMA that are not contiguous to the urban core are called urban fringe. Rural areas in the CMA are called rural fringe.

All CMAs are subdivided into census tracts.

In the 2006 Census, the CMA of Ottawa-Gatineau (Ontario-Quebec) crossed provincial boundaries. When the geographic level selected is all of Canada, the totals include the CMA on both sides of the provincial border. If a province has been selected, only the part of the CMA in the province chosen is included in the totals.

Census subdivision (CSD)

Census subdivision (CSD) is the general term applying to municipalities (as determined by provincial legislation) or their equivalent (for example, Indian reserves, Indian settlements and unorganized territories).

In Newfoundland and Labrador, Nova Scotia, and British Columbia, the term also describes geographic areas that have been created by Statistics Canada in cooperation with the provinces as equivalents for municipalities for the dissemination of statistical data.

Cohort

Represents a group of persons who have experienced a specific demographic event during a given year. In the case of births, persons born within a specified year are referred to as a generation.

Cohort component approach

This is the method used to produce estimates by age and sex, whereby the population is aged from year to year, and the components of demographic change are organised according to age and sex cohorts. Data required for this method

include demographic events such as deaths, immigration, emigration, etc. that can be directly linked to persons belonging to the same birth and sex cohorts.

Common-law union

Union consisting of two people of opposite sex or of the same sex who live together as a couple, without being legally married.

Component method

A method of generating population estimates which uses the components of demographic change and a base population as the input.

Components of demographic growth

Any of the classes of events generating population movement variations. Births, deaths, migration, marriages, divorces and new widowhoods are the components responsible for the variations since they alter either the total population or the age, sex and marital status distribution of the population.

Economic region (ER)

Refers to a group of complete census divisions (with one exception in Ontario) created as a standard geographic unit for analysis of regional economic activity.

Within the province of Quebec, economic regions ("regions administratives") are designated by law. In all other provinces or territories, economic regions are created by agreement between Statistics Canada and the provinces or territories concerned. Prince Edward Island and the three territories each consist of one economic region. In Ontario, there is one exception where the economic region boundary does not respect census division boundaries: the census division of Halton is split between the ER of Hamilton–Niagara Peninsula and the ER of Toronto. Prince Edward Island and the three territories each consist of one economic region.

Emigrant

Canadian citizen or immigrant who has left Canada to establish a permanent residence in another country.

Emigration

Represents departures from Canada of Canadian citizens or immigrants to another country, involving a change in usual place of residence. Emigration may be either temporary or permanent. Where the term 'emigration' is used alone in this

manual, it references permanent emigration which involves severing residential ties with Canada and acquiring permanent residency in another country.

Error of closure

Difference between the postcensal estimate for the same date as the census and the results of the census adjusted for census net undercoverage (including adjustment for incompletely enumerated Indian reserves).

Household

Refers to a person or a group of persons (other than foreign residents) who occupy the same dwelling and do not have a usual place of residence elsewhere in Canada. It may consist of a family group (census family) with or without other persons, of two or more families sharing a dwelling, of a group of unrelated persons, or of one person living alone.

Immigrant

Within the framework of this publication, the term immigrant refers to immigrant or permanent resident. An immigrant is a person who does not have Canadian citizenship at birth but was granted the right by immigration authorities to live in Canada on a permanent basis. Children born in other countries to parents who are Canadian citizens that reside temporarily in another country are not included in the category.

Immigration

Immigration represents entries into Canada of immigrants from another country, involving a change in usual place of residence.

Internal migration

Movements of persons within Canada's geographical boundaries, involving a change in usual place of residence. Internal migration denotes movement from one province or territory to another (i.e., interprovincial migration) or movements from some other smaller defined geographical unit to another (i.e., intraprovincial migration).

International migration

International migration represents movement of population between Canada and a foreign country which involves a change of the usual place of residence. A distinction is made with regard to immigrants, emigrants, returning emigrants, net temporary emigration, and net non-permanent residents.

Interprovincial migration

Interprovincial migration represents movement from one province or territory to another involving a permanent change in residence. A person who takes up residence in another province or territory is an out-migrant with reference to the province or territory of origin and an in-migrant with reference to the province or territory of destination.

Intraprovincial migration or subprovincial migration

Movement from one region to another within the same province or territory involving a permanent change of residence. A person who takes up residence in another region is an out-migrant with reference to the region of origin and an in-migrant with reference to the region of destination.

Legal marital status

Indicates the legal conjugal arrangement of a person. Estimates are presented in the following categories: single, legally married (including persons who are separated), widowed or divorced.

Single: Persons who have never been married, or persons whose marriage has been annulled and have not remarried.

Legally Married: Persons whose spouse is living, unless a divorce has been obtained. Persons separated are also included in this category.

Widowed: Persons who have lost their spouse through death and who have not remarried.

Divorced: Persons who have obtained a divorce and who have not remarried.

Marital status

Indicates the conjugal arrangement of a person. Estimates are presented in the following categories: single, married (including persons living common-law and persons who are separated), widowed or divorced.

Single: Persons who have never been married, or persons whose marriage has been annulled and who have not remarried, and do not live common-law.

Married (including persons living common-law and persons who are separated): Persons whose spouse is living, unless a divorce has been obtained.

Widowed: Persons who have lost their spouse through death and who have not remarried, and who do not live common-law.

Divorced: Persons who have obtained a divorce and who have not remarried, and who do not live common-law.

Natural increase

Variation of the population size over a given period as a result of the difference between the numbers of births and deaths.

Net internal migration

Sum of net intraprovincial and net interprovincial migration.

Net international migration

Net international migration is obtained according to the following formula: Immigrants + returning emigrants + net non-permanent residents – (emigrants + net temporary emigration).

Net interprovincial migration

Net interprovincial migration represents the difference between in-migrants and out-migrants for a given province or territory.

Net intraprovincial migration

Net intraprovincial migration represents the difference between in-migrants and out-migrants in a given intraprovincial region. A region can be defined as a census division, an economic region or a census metropolitan area.

Net non-permanent residents

Net non-permanent residents represent the variation in the number of non-permanent residents between two dates.

Net temporary emigration

Net temporary emigration represents the variation in the number of temporary emigrants between two dates. Temporary emigration includes Canadian citizens and immigrants living temporarily abroad who have not maintained a usual place of residence in Canada.

Non-permanent residents

Non-permanent residents are persons who are lawfully in Canada on a temporary basis under the authority of a temporary resident permit, along with members of their family living with them. Non-permanent residents include foreign workers, foreign students, the humanitarian population and other temporary residents. The humanitarian population includes refugee claimants and temporary residents who are allowed to remain in Canada on humanitarian grounds and are not categorized as either foreign workers or foreign students. For Demography Division, the terms non-permanent resident and temporary resident are equivalent.

Permanent resident

See immigrant.

Population

Estimated population and population according to the census are both defined as being the number of Canadians whose usual place of residence is in that area, regardless of where they happened to be on Census Day. Also included are any Canadians staying in a dwelling in that area on Census Day and having no usual place of residence elsewhere in Canada, as well as those considered non-permanent residents.

Population estimate

Postcensal: Population estimate produced by using data from the most recent available census adjusted for census net undercoverage (including adjustment for incompletely enumerated Indian reserves) and estimate of the components of demographic growth since that last census. This estimate can be preliminary, updated or final.

Intercensal: Population estimate derived by using postcensal estimates and data adjusted for census net undercoverage (including adjustment for incompletely enumerated Indian reserves) of censuses preceding and following the year in question.

Population growth or total growth

Variation of population size between two dates. It can also be obtained by summing the natural increase, total net migration and if possible, subtract residual deviation. It can be positive or negative.

Rate

Refers to the ratio of the number of events estimated in a period (from time period t to t+1, usually a year is the period) to the average populations at the beginning and the end of the period. In this regard, we calculate births, deaths, immigration rates. Generally, the rate is expressed in per 1,000.

Demographic growth rate or population growth: Ratio of population growth between the year t and t+1, to the average population of both these years. The rate is generally expressed in per 1,000.

Census net undercoverage rate: Difference between undercoverage rate and overcoverage rate.

Census overcoverage rate: The ratio of the numbers of persons or families who should not have been counted in the census or who were counted more than once to the total number of persons that should have been enumerated in the census. Generally, the rate is expressed in percentage.

Census undercoverage rate: The ratio of the estimated number of persons or families not enumerated in the census (who were intended to have been enumerated) to the total number of persons that should have been enumerated in the census. Generally, the rate is expressed in percentage.

Residual deviation

Difference between demographic population growths calculated using intercensal estimates of population between two dates and that obtained by the sum of the components for the same period. This deviation results from the distribution of the error of closure between years within the quinquennial period. This distribution is calculated by taking into account the number of days within each month.

Returning emigrant

Canadian citizen or immigrant having previously emigrated from Canada and subsequently returned to the country.

Reverse Record Check (RRC)

The Reverse Record Check is one of three studies which provide estimates of census coverage error, including independent estimates of census undercoverage and overcoverage. The RRC begins by working with a sample of all persons who were enumerated (or missed) in the previous census, along with all persons who were either born or entered into Canada over the intercensal period. With the exception of a very small sub-population of returning emigrants, the RRC

sampling frame includes all persons who could potentially be part of the census target universe.

Total net migration

Sum of net international and net internal migration.

Vital statistics

Vital Statistics includes all the demographic events (that is to say births, deaths, marriages and divorces) for which there are a legal requirement to inform the Provincial or Territorial Registrar's Office.

Year

Unless otherwise specified, the term "year" refers to the period beginning July 1 of a given year and ending June 30 of the following year.

References

Bédard, M. and M. Michalowski. 1997. *Advantages of the one year mobility variable for breaking down interprovincial migration by age, sex and marital status.* Demographic documents No. 7. Statistics Canada Catalogue no. 91-F0015MIE. Ottawa.

Dick, P. 1995. "Modeling Net Undercoverage in the 1991 Canadian Census." *Survey Methodology.* Vol. 21, no. 1, p. 45-54. Statistics Canada. Ottawa.

Fay, R.E. and R.A. Herriot. 1979. "Estimates of Income for Small Places: An Application of James-Stein Procedures to Census Data." *Journal of the American Statistical Association.* Vol. 81, p. 269-277.

Felligi, I.P. and Sunter, A.B. 1969. "A Theory of Record Linkage." *Journal of the American Statistical Association,* Vol. 64, p. 1,183-1,210.

Gambino, J. and Dick, P. 2000. "Small Area Estimation Practices at Statistics Canada." *Statistics in Transition: Journal of the Polish Statistical Association.* Vol. 4, p. 597-610.

Michalowski, M. 1999. *Canadian Residing Temporarily Abroad: Numbers, characteristics and Estimation Methods.* Working paper. Statistics Canada. Ottawa.

Ramsay, T. 2000. *Small Domain Estimation Based on Nonparametric Regression,* Working paper. Statistics Canada. Ottawa.

Shryock, H.S., J.S. Siegel and Associates. 1976. *Studies in Population: The Methods and Materials of Demography.* Washington: Academic Press, Inc.

Statistics Canada. 2011a. *Quarterly Demographic Estimates January to March 2011.* Statistics Canada Catalogue no. 91-002-X. Ottawa.

--. 2011b. *Annual Demographic Estimates: Subprovincial Areas 2005 to 2010.* Statistics Canada Catalogue no. 91-214-X. Ottawa.

--. 2010a. *Annual Demographic Estimates: Canada, Provinces and Territories 2010.* Statistics Canada Catalogue no. 91-215-X. Ottawa.

--. 2010b. *Population Projections for Canada, Provinces and Territories 2009 to 2036.* Statistics Canada Catalogue no. 91-520-X. Ottawa.

--. 2010c. *2006 Census Technical Report: Coverage.* Statistics Canada Catalogue no. 92-567-X. Ottawa.
http://www12.statcan.gc.ca/census-recensement/2006/ref/rp-guides/rp/coverage-couverture/cov-couv_af-fdr-eng.cfm
(accessed June 29, 2011)

--. 2010d. *Migration Estimates – User's Guide*. Income Statistics Division, Statistics Canada. Ottawa.

--. 2007a. *2006 Census Dictionary*. Statistics Canada Catalogue no. 92-566-XWE. Ottawa.
http://www12.statcan.gc.ca/census-recensement/2006/ref/dict/index-eng.cfm
(accessed July 5, 2011).

--. 2007b . *Population and Family Estimation Methods at Statistics Canada.* Statistics Canada Catalogue no. 91-528-XIE. Ottawa.

--. 2003. *Population and Family Estimation Methods at Statistics Canada.* Statistics Canada Catalogue no. 91-528-XIE. Ottawa.

--. 1987. *Population Estimation Methods, Canada.* Statistics Canada Catalogue no. 91-528E. Ottawa.

United Nations. 1998. *Recommendations on Statistics of International Migration – Revision 1.* New York. Department of Economic and Social Affairs, Statistics Division. Statistical Papers, Series M, No. 58. Rev. 1.

Wilkinson, P. 2004. *Estimates of Internal Migration Based on New and Old Methods for Combined Annual Periods 1996-1997 to 2000-2001.* Demographic documents No. 6. Statistics Canada Catalogue no. 91-F0015MIE. Ottawa.